

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/72, G01N 33/50, 33/566	A1	(11) International Publication Number: WO 00/22129 (43) International Publication Date: 20 April 2000 (20.04.00)
(21) International Application Number: PCT/US99/23938 (22) International Filing Date: 12 October 1999 (12.10.99) (30) Priority Data: 09/170,496 13 October 1998 (13.10.98) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 09/170,496 (CIP) Filed on 13 October 1998 (13.10.98) (71) Applicant (for all designated States except US): ARENA PHARMACEUTICALS, INC. [US/US]; 6166 Nancy Ridge Drive, San Diego, CA 92121 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): BEHAN, Dominic, P. [GB/US]; 11472 Roxboro Court, San Diego, CA 92131 (US). CHALMERS, Derek, T. [GB/US]; 347 Longden Lane, Solana Beach, CA 92075 (US). LIAW, Chen, W. [US/US]; 7668 Salix Place, San Diego, CA 92129 (US).		(74) Agents: MILLER, Suzanne, E. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris LLP, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US). (81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: NON-ENDOGENOUS, CONSTITUTIVELY ACTIVATED HUMAN G PROTEIN-COUPLED RECEPTORS (57) Abstract Disclosed herein are constitutively activated, non-endogenous versions of endogenous human G protein-coupled receptors comprising (a) the following amino acid sequence region (C-terminus to N-terminus orientation) and/or (b) the following nucleic acid sequence region (3' to 5' orientation) transversing the transmembrane-6 (TM6) and intracellular loop-3 (IC3) regions of the GPCR: (a) P ¹ AA ₁₅ X and/or (b) P ^{codon} (AA-codon) ₁₅ X _{codon} , respectively. In a most preferred embodiment, P ¹ and P ^{codon} are endogenous proline and an endogenous nucleic acid encoding region encoding proline, respectively, located within TM6 of the non-endogenous GPCR; AA ₁₅ and (AA-codon) ₁₅ are 15 endogenous amino acid residues and 15 codons encoding endogenous amino acid residues, respectively; and X and X _{codon} are non-endogenous lysine and a non-endogenous nucleic acid encoding region encoding lysine, respectively, located within IC3 of the non-endogenous GPCR. Because it is most preferred that the non-endogenous human GPCRs which incorporate these mutations are incorporated into mammalian cells and utilized for the screening of the candidate compounds, the non-endogenous human GPCR incorporating the mutation need not be purified and isolated <i>per se</i> (i.e., these are incorporated within the cellular membrane of a mammalian cell), although such purified and isolated non-endogenous human GPCRs are well within the purview of this disclosure.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

FIGURE 3 L

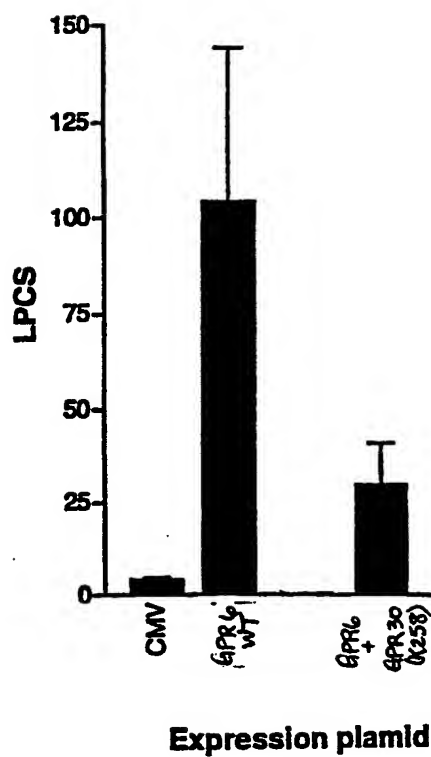


FIGURE 4

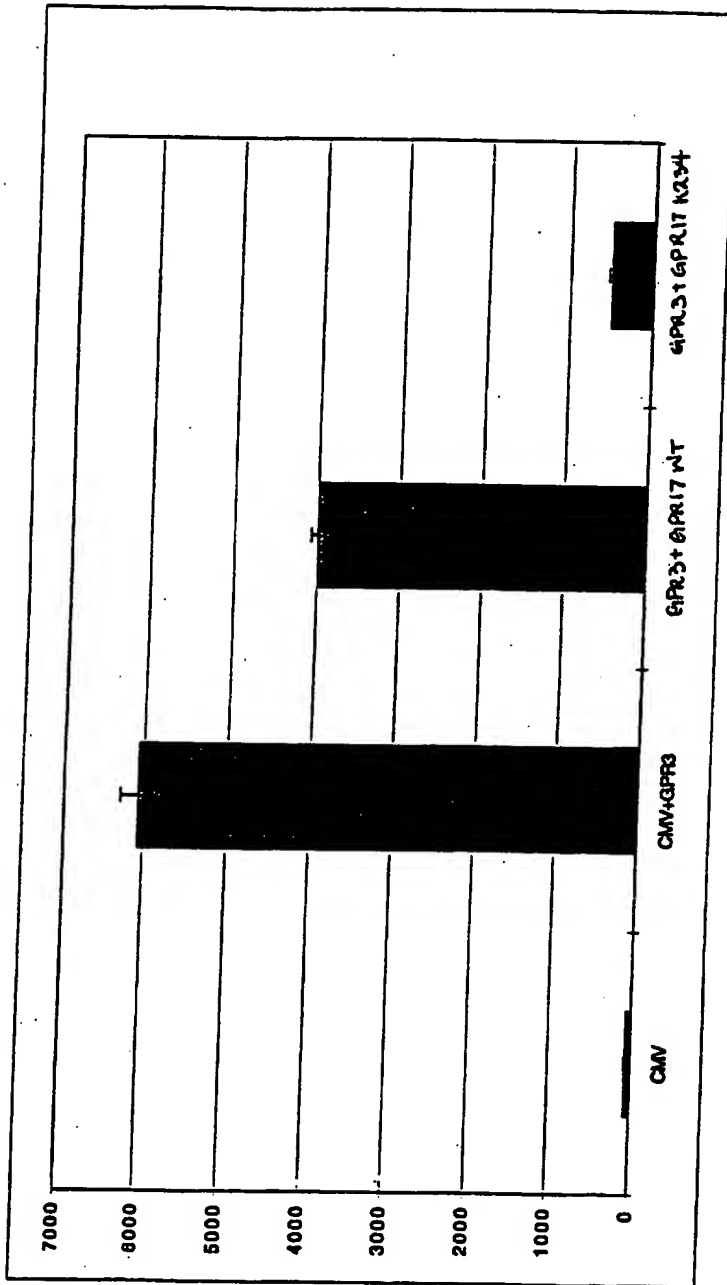


FIGURE 5

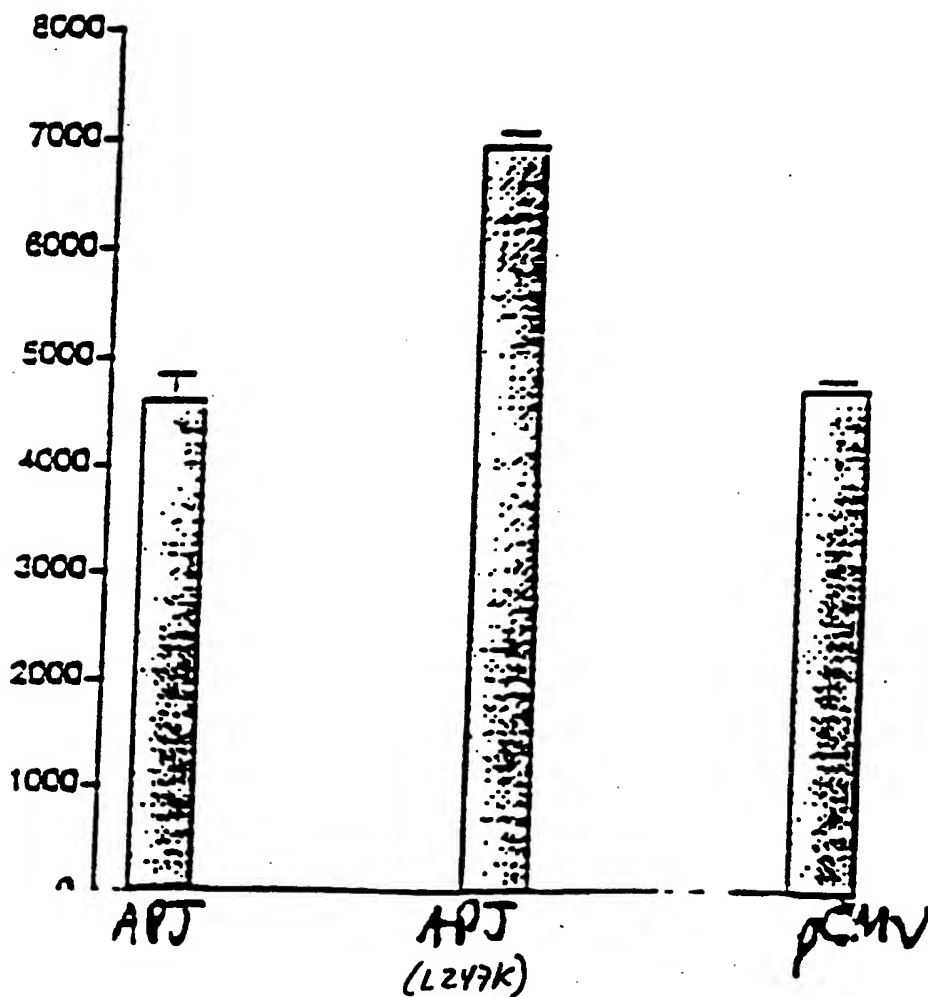


FIGURE 6

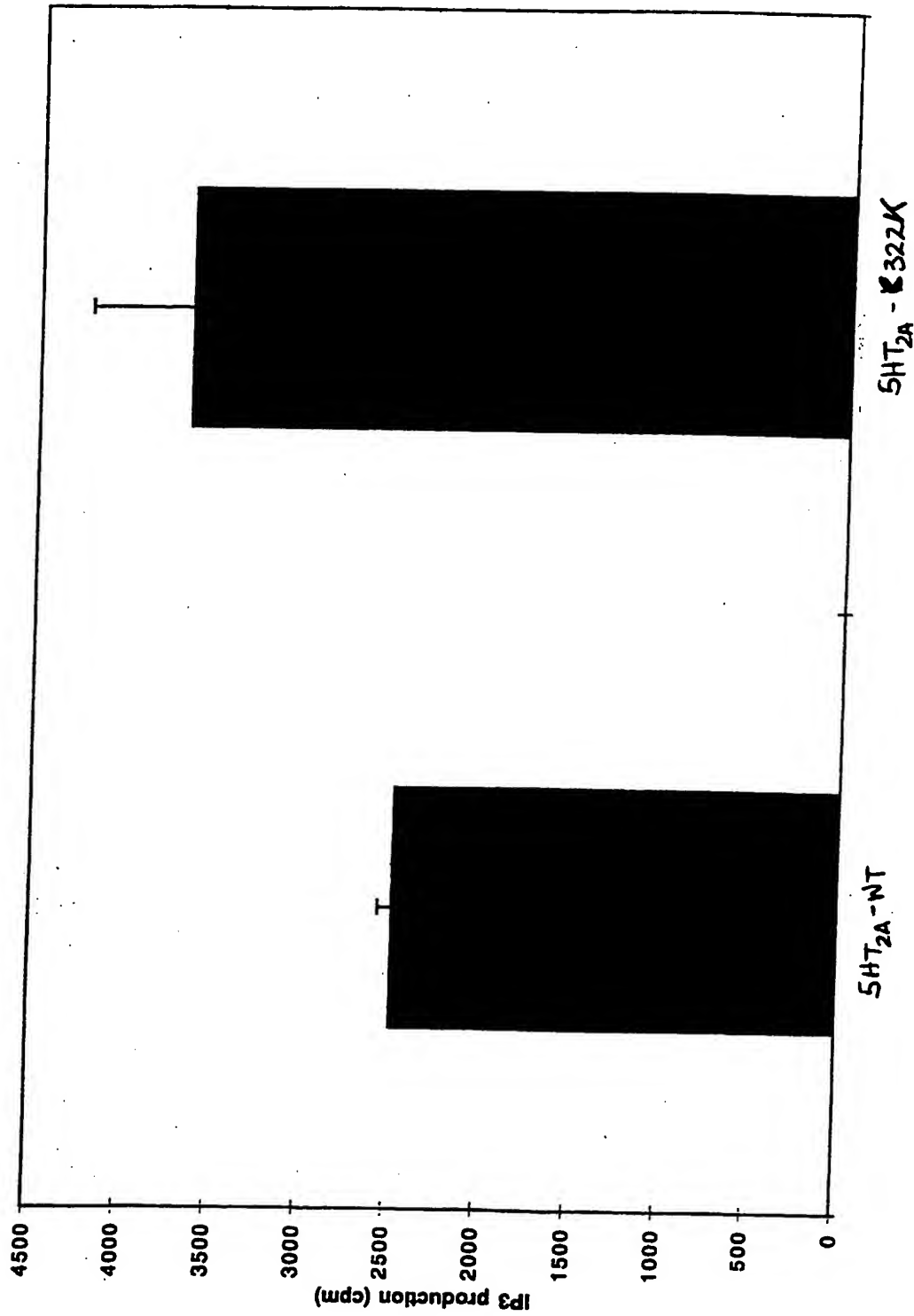


FIGURE 7

FIGURE 8A

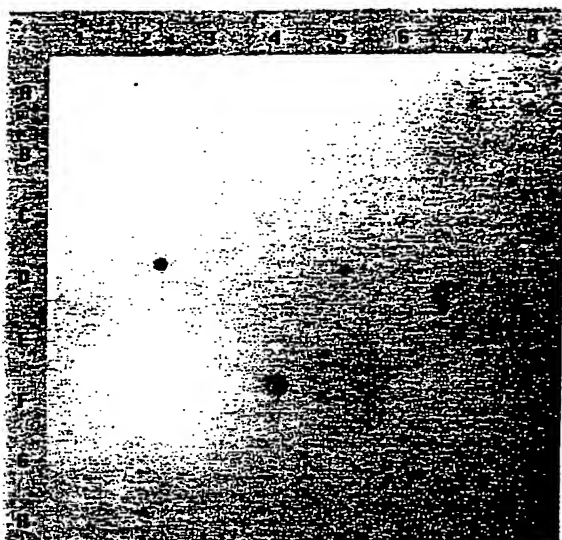


FIGURE 8B

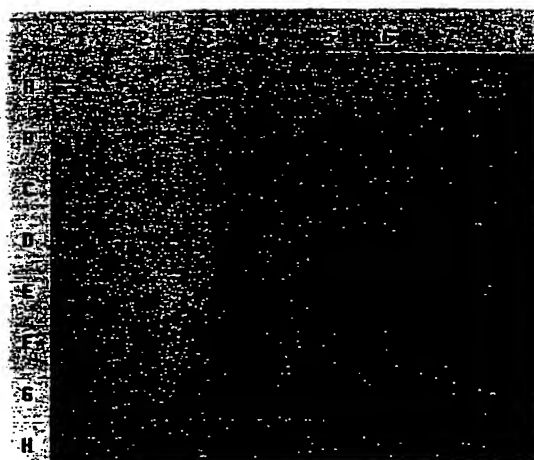
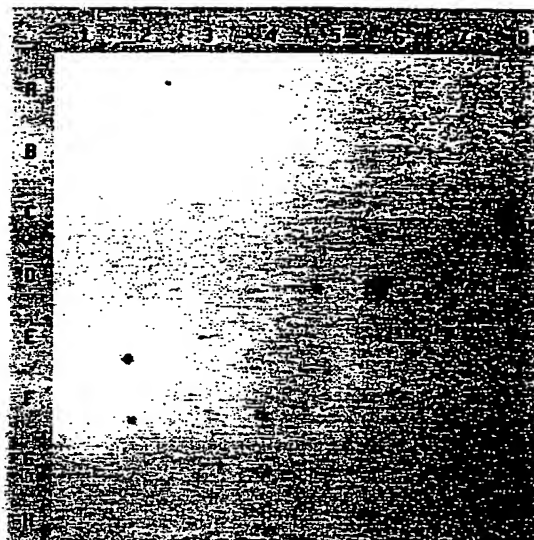


FIGURE 8C

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Behan, Dominic P.
Chalmers, Derek T.
Liaw, Chen W.
- (ii) TITLE OF INVENTION: Non-Endogenous, Constitutively
Activated Human G Protein-Coupled
Orphan Receptors
- 10 (iii) NUMBER OF SEQUENCES: 280
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Arena Pharmaceuticals, Inc.
15 (B) STREET: 6166 Nancy Ridge Drive
(C) CITY: San Diego
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92122
- (v) COMPUTER READABLE FORM:
- 20 (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
- 25 (A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Burgoon, Richard P.
(B) REGISTRATION NUMBER: 34,787
- 30 (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (619)453-7200
(B) TELEFAX: (619)453-7210

(2) INFORMATION FOR SEQ ID NO:1:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTTGAAA ACTATTCCTA TGACCTAGAC 60

TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120
 TCCCTGGTGT TATATTGTTT GGCTTTTGTT CTGGGAATTC CAGGAAATGC CATCGTCATT 180
 TGGTTCACGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240
 ATTGCGGATT TCATTTTTCT TCTCTTTCTG CCCCTGTACA TCTCCTATGT GGCCATGAAT 300
 5 TTCCACTGGC CCTTTGGCAT CTGGCTGTGC AAAGCCAATT CCTTCACTGC CCAGTTGAAC 360
 ATGTTTGCCA GTGTTTTTTT CCTGACAGTG ATCAGCCTGG ACCACTATAT CCACTTGATC 420
 CATCCTGTCT TATCTCATCG GCATCGAACC CTCAAGAACT CTCTGATTGT CATTATATTC 480
 ATCTGGCTTT TGGCTTCTCT AATTGGCGGT CCTGCCCTGT ACTTCCGGGA CACTGTGGAG 540
 TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTCACTTTG 600
 10 ATCAGGCACC ATGTTCTGAC TTGGGTGAAA TTTATCATTG GCTATCTCTT CCCTTTGCTA 660
 ACAATGAGTA TTTGCTACTT GTGTCTCATC TTCAAGGTGA AGAAGCGAAC AGTCCTGATC 720
 TCCAGTAGGC ATTTCTGGAC AATTCTGGTT GTGGTTGTGG CCTTTGTGGT TTGCTGGACT 780
 CCTTATCACC TGTTTAGCAT TTGGGAGCTC ACCATTACACC ACAATAGCTA TTCCCACCAT 840
 GTGATGCAGG CTGGAATCCC CCTCTCCACT GGTTTGGCAT TCCTCAATAG TTGCTTGAAC 900
 15 CCCATCCTTT ATGTCCTAAT TAGTAAGAAG TTCCAAGCTC GCTTCCGGTC CTCAGTTGCT 960
 GAGATACTCA AGTACACACT GTGGGAAGTC AGCTGTTCTG GCACAGTGAG TGAACAGCTC 1020
 AGGAACTCAG AAACCAAGAA TCTGTGTCTC CTGGAAACAG CTCAATAA 1068

(3) INFORMATION FOR SEQ ID NO:2:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 355 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Asp Leu Glu Glu Thr Leu Phe Glu Glu Phe Glu Asn Tyr Ser
 1 5 10 15

Tyr Asp Leu Asp Tyr Tyr Ser Leu Glu Ser Asp Leu Glu Glu Lys Val
 20 25 30

30 Gln Leu Gly Val Val His Trp Val Ser Leu Val Leu Tyr Cys Leu Ala
 35 40 45

Phe Val Leu Gly Ile Pro Gly Asn Ala Ile Val Ile Trp Phe Thr Gly
 50 55 60

Leu Lys Trp Lys Lys Thr Val Thr Thr Leu Trp Phe Leu Asn Leu Ala
 65 70 75 80

5 Ile Ala Asp Phe Ile Phe Leu Leu Phe Leu Pro Leu Tyr Ile Ser Tyr
 85 90 95

Val Ala Met Asn Phe His Trp Pro Phe Gly Ile Trp Leu Cys Lys Ala
 100 105 110

10 Asn Ser Phe Thr Ala Gln Leu Asn Met Phe Ala Ser Val Phe Phe Leu
 115 120 125

Thr Val Ile Ser Leu Asp His Tyr Ile His Leu Ile His Pro Val Leu
 130 135 140

Ser His Arg His Arg Thr Leu Lys Asn Ser Leu Ile Val Ile Ile Phe
 145 150 155 160

15 Ile Trp Leu Leu Ala Ser Leu Ile Gly Gly Pro Ala Leu Tyr Phe Arg
 165 170 175

Asp Thr Val Glu Phe Asn Asn His Thr Leu Cys Tyr Asn Asn Phe Gln
 180 185 190

20 Lys His Asp Pro Asp Leu Thr Leu Ile Arg His His Val Leu Thr Trp
 195 200 205

Val Lys Phe Ile Ile Gly Tyr Leu Phe Pro Leu Leu Thr Met Ser Ile
 210 215 220

Cys Tyr Leu Cys Leu Ile Phe Lys Val Lys Lys Arg Thr Val Leu Ile
 225 230 235 240

25 Ser Ser Arg His Phe Trp Thr Ile Leu Val Val Val Val Ala Phe Val
 245 250 255

Val Cys Trp Thr Pro Tyr His Leu Phe Ser Ile Trp Glu Leu Thr Ile
 260 265 270

30 His His Asn Ser Tyr Ser His His Val Met Gln Ala Gly Ile Pro Leu
 275 280 285

Ser Thr Gly Leu Ala Phe Leu Asn Ser Cys Leu Asn Pro Ile Leu Tyr
 290 295 300

Val Leu Ile Ser Lys Lys Phe Gln Ala Arg Phe Arg Ser Ser Val Ala
 305 310 315 320

35 Glu Ile Leu Lys Tyr Thr Leu Trp Glu Val Ser Cys Ser Gly Thr Val
 325 330 335

Ser Glu Gln Leu Arg Asn Ser Glu Thr Lys Asn Leu Cys Leu Leu Glu

340

345

350

Thr Ala Gln
355

(4) INFORMATION FOR SEQ ID NO:3:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1089 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGACTCGC GCGTGGACCA CCTCTTTCCG 60
CCATCCCTCT ACATCTTTGT CATCGGCGTG GGGCTGCCCA CCAACTGCCT GGCTCTGTGG 120
GCGGCCTACC GCCAGGTGCA ACAGCGCAAC GAGCTGGGCG TCTACCTGAT GAACCTCAGC 180
15 ATCGCCGACC TGCTGTACAT CTGCACGCTG CCGCTGTGGG TGGACTACTT CCTGCACCAC 240
GACAACTGGA TCCACGGCCC CGGGTCCTGC AAGCTCTTTG GGTTCATCTT CTACACCAAT 300
ATCTACATCA GCATCGCCTT CCTGTGCTGC ATCTCGGTGG ACCGCTACCT GGCTGTGGCC 360
CACCCACTCC GCTTCGCCCCG CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGCTCCGTG 420
GTCTGGGCCA CGGAGCTGGG CGCCAACTCG GCGCCCCTGT TCCATGACGA GCTCTTCCGA 480
20 GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCCA TGGAAGGCTG GGTGGCCTGG 540
ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTTCCCGT GGGCGCTCAT GCTGCTGTCTG 600
TACCGGGGCA TCCTGCGGGC CGTGCAGGGC AGCGTGTTCA CCGAGCGCCA GGAGAAGGCC 660
AAGATCAAGC GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TGCGCCCTAT 720
CACGTGCTCT TGCTGTCCCG CAGCGCCATC TACCTGGGCC GCCCCTGGGA CTGCGGCTTC 780
25 GAGGAGCGCG TCTTTTCTGC ATACCACAGC TCACTGGCTT TCACCAGCCT CAACTGTGTG 840
GCGGACCCCA TCCTCTACTG CCTGGTCAAC GAGGGCGCCC GCAGCGATGT GGCCAAGGCC 900
CTGCACAACC TGCTCCGCTT TCTGGCCAGC GACAAGCCCC AGGAGATGGC CAATGCCCTCG 960
CTCACCTTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC 1020
AGCTGGGCGG CCACTCCGCC TTCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA 1080
30 GCACAATGA 1089

(5) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 362 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

10 Met Gly Asn His Thr Trp Glu Gly Cys His Val Asp Ser Arg Val Asp
 1 5 10 15
 His Leu Phe Pro Pro Ser Leu Tyr Ile Phe Val Ile Gly Val Gly Leu
 20 25 30
 Pro Thr Asn Cys Leu Ala Leu Trp Ala Ala Tyr Arg Gln Val Gln Gln
 35 40 45
 15 Arg Asn Glu Leu Gly Val Tyr Leu Met Asn Leu Ser Ile Ala Asp Leu
 50 55 60
 Leu Tyr Ile Cys Thr Leu Pro Leu Trp Val Asp Tyr Phe Leu His His
 65 70 75 80
 20 Asp Asn Trp Ile His Gly Pro Gly Ser Cys Lys Leu Phe Gly Phe Ile
 85 90 95
 Phe Tyr Thr Asn Ile Tyr Ile Ser Ile Ala Phe Leu Cys Cys Ile Ser
 100 105 110
 Val Asp Arg Tyr Leu Ala Val Ala His Pro Leu Arg Phe Ala Arg Leu
 115 120 125
 25 Arg Arg Val Lys Thr Ala Val Ala Val Ser Ser Val Val Trp Ala Thr
 130 135 140
 Glu Leu Gly Ala Asn Ser Ala Pro Leu Phe His Asp Glu Leu Phe Arg
 145 150 155 160
 30 Asp Arg Tyr Asn His Thr Phe Cys Phe Glu Lys Phe Pro Met Glu Gly
 165 170 175
 Trp Val Ala Trp Met Asn Leu Tyr Arg Val Phe Val Gly Phe Leu Phe
 180 185 190
 Pro Trp Ala Leu Met Leu Leu Ser Tyr Arg Gly Ile Leu Arg Ala Val
 195 200 205
 35 Arg Gly Ser Val Ser Thr Glu Arg Gln Glu Lys Ala Lys Ile Lys Arg
 210 215 220
 Leu Ala Leu Ser Leu Ile Ala Ile Val Leu Val Cys Phe Ala Pro Tyr

6

225 230 235 240
 His Val Leu Leu Leu Ser Arg Ser Ala Ile Tyr Leu Gly Arg Pro Trp
 245 250 255
 5 Asp Cys Gly Phe Glu Glu Arg Val Phe Ser Ala Tyr His Ser Ser Leu
 260 265 270
 Ala Phe Thr Ser Leu Asn Cys Val Ala Asp Pro Ile Leu Tyr Cys Leu
 275 280 285
 Val Asn Glu Gly Ala Arg Ser Asp Val Ala Lys Ala Leu His Asn Leu
 290 295 300
 10 Leu Arg Phe Leu Ala Ser Asp Lys Pro Gln Glu Met Ala Asn Ala Ser
 305 310 315 320
 Leu Thr Leu Glu Thr Pro Leu Thr Ser Lys Arg Asn Ser Thr Ala Lys
 325 330 335
 15 Ala Met Thr Gly Ser Trp Ala Ala Thr Pro Pro Ser Gln Gly Asp Gln
 340 345 350
 Val Gln Leu Lys Met Leu Pro Pro Ala Gln
 355 360

(6) INFORMATION FOR SEQ ID NO:5:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TATGAATTCA GATGCTCTAA ACGTCCCTGC

30

(7) INFORMATION FOR SEQ ID NO:6:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

35 TCCGGATCCA CCTGCACCTG CGCCTGCACC

30

(8) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1002 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG 60
CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTCGCCA CCACTGTCCT GTACTGCCTG 120
10 GTGTTTCTCC TCAGCCTAGT GGGCAACAGC CTGGTCTGT GGGTCCTGGT GAAGTATGAG 180
AGCCTGGAGT CCCTCACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTC 240
GCCTGCTTGT TGCCTGTGTG GATCTCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC 300
CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCCTG 360
ACCATCATGA CCATCCACCG CTACCTGTCT GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC 420
15 CCCACCCTCC GCTGCCGGGT GCTGGTGACC ATGGCTGTGT GGGTAGCCAG CATCCTGTCC 480
TCCATCCTCG ACACCATCTT CCACAAGGTG CTTTCTTCGG GCTGTGATTA TTCCGAACTC 540
ACGTGGTACC TCACCTCCGT CTACCAGCAC AACCTCTTCT TCCTGCTGTC CCTGGGGATT 600
ATCCTGTTCT GCTACGTGGA GATCCTCAGG ACCCTGTTCC GCTCACGCTC CAAGCGGCGC 660
CACCGCACGG TCAAGCTCAT CTTGCCATC GTGGTGGCCT ACTTCCTCAG CTGGGGTCCC 720
20 TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCGGACCC AGATCATCCG GAGCTGCGAG 780
GCCAAACAGC AGCTAGAATA CGCCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCACTGC 840
TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT 900
GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCCAC 960
TCCCCTGGTG CCTTCGCCTA TGAGGGCGCC TCCTTCTACT GA 1002

25 (9) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Glu Ser Ser Gly Asn Pro Glu Ser Thr Thr Phe Phe Tyr Tyr Asp
 1 5 10 15
 Leu Gln Ser Gln Pro Cys Glu Asn Gln Ala Trp Val Phe Ala Thr Leu
 20 25 30
 5 Ala Thr Thr Val Leu Tyr Cys Leu Val Phe Leu Leu Ser Leu Val Gly
 35 40 45
 Asn Ser Leu Val Leu Trp Val Leu Val Lys Tyr Glu Ser Leu Glu Ser
 50 55 60
 10 Leu Thr Asn Ile Phe Ile Leu Asn Leu Cys Leu Ser Asp Leu Val Phe
 65 70 75 80
 Ala Cys Leu Leu Pro Val Trp Ile Ser Pro Tyr His Trp Gly Trp Val
 85 90 95
 Leu Gly Asp Phe Leu Cys Lys Leu Leu Asn Met Ile Phe Ser Ile Ser
 100 105 110
 15 Leu Tyr Ser Ser Ile Phe Phe Leu Thr Ile Met Thr Ile His Arg Tyr
 115 120 125
 Leu Ser Val Val Ser Pro Leu Ser Thr Leu Arg Val Pro Thr Leu Arg
 130 135 140
 20 Cys Arg Val Leu Val Thr Met Ala Val Trp Val Ala Ser Ile Leu Ser
 145 150 155 160
 Ser Ile Leu Asp Thr Ile Phe His Lys Val Leu Ser Ser Gly Cys Asp
 165 170 175
 Tyr Ser Glu Leu Thr Trp Tyr Leu Thr Ser Val Tyr Gln His Asn Leu
 180 185 190
 25 Phe Phe Leu Leu Ser Leu Gly Ile Ile Leu Phe Cys Tyr Val Glu Ile
 195 200 205
 Leu Arg Thr Leu Phe Arg Ser Arg Ser Lys Arg Arg His Arg Thr Val
 210 215 220
 30 Lys Leu Ile Phe Ala Ile Val Val Ala Tyr Phe Leu Ser Trp Gly Pro
 225 230 235 240
 Tyr Asn Phe Thr Leu Phe Leu Gln Thr Leu Phe Arg Thr Gln Ile Ile
 245 250 255
 Arg Ser Cys Glu Ala Lys Gln Gln Leu Glu Tyr Ala Leu Leu Ile Cys
 260 265 270
 35 Arg Asn Leu Ala Phe Ser His Cys Cys Phe Asn Pro Val Leu Tyr Val
 275 280 285

Phe Val Gly Val Lys Phe Arg Thr His Leu Lys His Val Leu Arg Gln
290 295 300

Phe Trp Phe Cys Arg Leu Gln Ala Pro Ser Pro Ala Ser Ile Pro His
305 310 315 320

5 Ser Pro Gly Ala Phe Ala Tyr Glu Gly Ala Ser Phe Tyr
325 330

(10) INFORMATION FOR SEQ ID NO:9:

- 10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

15 GCAAGCTTGG GGGACGCCAG GTCGCCGGCT 30

(11) INFORMATION FOR SEQ ID NO:10:

- 20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGGATCCGG ACGCTGGGGG AGTCAGGCTG C 31

25 (12) INFORMATION FOR SEQ ID NO:11:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 987 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGGACAACG CCTCGTTCTC GGAGCCCTGG CCCGCCAACG CATCGGGCCC GGACCCGGCG 60

CTGAGCTGCT CCAACGCGTC GACTCTGGCG CCGCTGCCGG CGCCGCTGGC GGTGGCTGTA 120

35 CCAAGTTGTCT ACGCGGTGAT CTGCGCCGTG GGTCTGGCGG GCAACTCCGC CGTGCTGTAC 180

GTGTTGCTGC GGGCGCCCCG CATGAAGACC GTCACCAACC TGTTCATCCT CAACCTGGCC 240
 ATCGCCGACG AGCTCTTCAC GCTGGTGCTG CCCATCAACA TCGCCGACTT CCTGTGCGG 300
 CAGTGGCCCT TCGGGGAGCT CATGTGCAAG CTCATCGTGG CTATCGACCA GTACAACACC 360
 TTCTCCAGCC TCTACTTCCT CACCGTCATG AGCGCCGACC GCTACCTGGT GGTGTTGGCC 420
 5 ACTGCGGAGT CGCGCCGGGT GGCCGGCCGC ACCTACAGCG CCGCGCGCGC GGTGAGCCTG 480
 GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
 GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
 CGCGCGAGCC GCCTCTACAC GCTCGTGCTG GGCTTCGCCA TCCCCGTGTC CACCATCTGT 660
 GTCCTCTATA CCACCCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
 10 GCCCTGGAGC GCGCCAAGAA GCGGGTGACC TTCCTGGTGG TGGCAATCCT GCGGGTGTGC 780
 CTCCTCTGCT GGACGCCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
 CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATCACCA GCCTGACGTA CGCCAACAGC 900
 TGCCTCAACC CCTTCCTCTA CGCCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
 CTGATAACTT GCCGCGCGGC AGCCTGA 987

15 (13) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 328 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - 20 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Asp Asn Ala Ser Phe Ser Glu Pro Trp Pro Ala Asn Ala Ser Gly
 1 5 10 15
 25 Pro Asp Pro Ala Leu Ser Cys Ser Asn Ala Ser Thr Leu Ala Pro Leu
 20 25 30
 Pro Ala Pro Leu Ala Val Ala Val Pro Val Val Tyr Ala Val Ile Cys
 35 40 45
 30 Ala Val Gly Leu Ala Gly Asn Ser Ala Val Leu Tyr Val Leu Leu Arg
 50 55 60
 Ala Pro Arg Met Lys Thr Val Thr Asn Leu Phe Ile Leu Asn Leu Ala
 65 70 75 80

11

Ile Ala Asp Glu Leu Phe Thr Leu Val Leu Pro Ile Asn Ile Ala Asp
85 90 95

Phe Leu Leu Arg Gln Trp Pro Phe Gly Glu Leu Met Cys Lys Leu Ile
100 105 110

5 Val Ala Ile Asp Gln Tyr Asn Thr Phe Ser Ser Leu Tyr Phe Leu Thr
115 120 125

Val Met Ser Ala Asp Arg Tyr Leu Val Val Leu Ala Thr Ala Glu Ser
130 135 140

10 Arg Arg Val Ala Gly Arg Thr Tyr Ser Ala Ala Arg Ala Val Ser Leu
145 150 155 160

Ala Val Trp Gly Ile Val Thr Leu Val Val Leu Pro Phe Ala Val Phe
165 170 175

Ala Arg Leu Asp Asp Glu Gln Gly Arg Arg Gln Cys Val Leu Val Phe
180 185 190

15 Pro Gln Pro Glu Ala Phe Trp Trp Arg Ala Ser Arg Leu Tyr Thr Leu
195 200 205

Val Leu Gly Phe Ala Ile Pro Val Ser Thr Ile Cys Val Leu Tyr Thr
210 215 220

20 Thr Leu Leu Cys Arg Leu His Ala Met Arg Leu Asp Ser His Ala Lys
225 230 235 240

Ala Leu Glu Arg Ala Lys Lys Arg Val Thr Phe Leu Val Val Ala Ile
245 250 255

Leu Ala Val Cys Leu Leu Cys Trp Thr Pro Tyr His Leu Ser Thr Val
260 265 270

25 Val Ala Leu Thr Thr Asp Leu Pro Gln Thr Pro Leu Val Ile Ala Ile
275 280 285

Ser Tyr Phe Ile Thr Ser Leu Thr Tyr Ala Asn Ser Cys Leu Asn Pro
290 295 300

30 Phe Leu Tyr Ala Phe Leu Asp Ala Ser Phe Arg Arg Asn Leu Arg Gln
305 310 315 320

Leu Ile Thr Cys Arg Ala Ala Ala
325

(14) INFORMATION FOR SEQ ID NO:13:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGAATTCGT CAACGGTCCC AGCTACAATG

30

(15) INFORMATION FOR SEQ ID NO:14:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATGGATCCCA GGCCCTTCAG CACCGCAATA T

31

(16) INFORMATION FOR SEQ ID NO:15:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1002 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGCAGGCCG CTGGGCACCC AGAGCCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG 60

ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA 120

CTGCCGTTCC TCTATGTGCT CCTGCCCGCC GTGTACTCCG GGATCTGTGC TGTGGGGCTG 180

ACTGGCAACA CGGCCGTCAT CCTTGTAATC CTAAGGGCGC CCAAGATGAA GACGGTGACC 240

25 AACGTGTTCA TCCTGAACCT GGCCGTCGCC GACGGGCTCT TCACGCTGGT ACTGCCCCGTC 300

AACATCGCGG AGCACCTGCT GCAGTACTGG CCCTTCGGGG AGCTGCTCTG CAAGCTGGTG 360

CTGGCCGTCG ACCACTACAA CATCTTCTCC AGCATCTACT TCCTAGCCGT GATGAGCGTG 420

GACCGATACC TGGTGGTGCT GGCCACCGTG AGGTCCCGCC ACATGCCCTG GCGCACCTAC 480

CGGGGGGCGA AGGTCGCCAG CCTGTGTGTC TGGCTGGGCG TCACGGTCCT GGTTCCTGCC 540

30 TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCAAGCTG TGGGCTGAGC 600

TTCCCGTGGC CCGAGCGGGT CTGGTTCAAG GCCAGCCGTG TCTACACTTT GGTCTGGGC 660

TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCCTGCG CAGGCTGCGG 720

GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGACCGTC 780
 CTGGTCCTCG TCGTGCTGGC CGTGTGCCTC CTCTGCTGGA CGCCCTTCCA CCTGGCCTCT 840
 GTCGTGGCCC TGACCACGGA CCTGCCCCAG ACCCCACTGG TCATCAGTAT GTCCTACGTC 900
 ATCACCAGCC TCACGTACGC CAACTCGTGC CTGAACCCCT TCCTCTACGC CTTTCTAGAT 960
 5 GACAACTTCC GGAAGAACTT CCGCAGCATA TTGCGGTGCT GA 1002

(17) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

15 Met Gln Ala Ala Gly His Pro Glu Pro Leu Asp Ser Arg Gly Ser Phe
 1 5 10 15
 Ser Leu Pro Thr Met Gly Ala Asn Val Ser Gln Asp Asn Gly Thr Gly
 20 25 30
 His Asn Ala Thr Phe Ser Glu Pro Leu Pro Phe Leu Tyr Val Leu Leu
 35 40 45
 20 Pro Ala Val Tyr Ser Gly Ile Cys Ala Val Gly Leu Thr Gly Asn Thr
 50 55 60
 Ala Val Ile Leu Val Ile Leu Arg Ala Pro Lys Met Lys Thr Val Thr
 65 70 75 80
 25 Asn Val Phe Ile Leu Asn Leu Ala Val Ala Asp Gly Leu Phe Thr Leu
 85 90 95
 Val Leu Pro Val Asn Ile Ala Glu His Leu Leu Gln Tyr Trp Pro Phe
 100 105 110
 Gly Glu Leu Leu Cys Lys Leu Val Leu Ala Val Asp His Tyr Asn Ile
 115 120 125
 30 Phe Ser Ser Ile Tyr Phe Leu Ala Val Met Ser Val Asp Arg Tyr Leu
 130 135 140
 Val Val Leu Ala Thr Val Arg Ser Arg His Met Pro Trp Arg Thr Tyr
 145 150 155 160
 35 Arg Gly Ala Lys Val Ala Ser Leu Cys Val Trp Leu Gly Val Thr Val
 165 170 175

14

Leu Val Leu Pro Phe Phe Ser Phe Ala Gly Val Tyr Ser Asn Glu Leu
 180 185 190
 Gln Val Pro Ser Cys Gly Leu Ser Phe Pro Trp Pro Glu Arg Val Trp
 195 200 205
 5 Phe Lys Ala Ser Arg Val Tyr Thr Leu Val Leu Gly Phe Val Leu Pro
 210 215 220
 Val Cys Thr Ile Cys Val Leu Tyr Thr Asp Leu Leu Arg Arg Leu Arg
 225 230 235 240
 10 Ala Val Arg Leu Arg Ser Gly Ala Lys Ala Leu Gly Lys Ala Arg Arg
 245 250 255
 Lys Val Thr Val Leu Val Leu Val Val Leu Ala Val Cys Leu Leu Cys
 260 265 270
 Trp Thr Pro Phe His Leu Ala Ser Val Val Ala Leu Thr Thr Asp Leu
 275 280 285
 15 Pro Gln Thr Pro Leu Val Ile Ser Met Ser Tyr Val Ile Thr Ser Leu
 290 295 300
 Thr Tyr Ala Asn Ser Cys Leu Asn Pro Phe Leu Tyr Ala Phe Leu Asp
 305 310 315 320
 20 Asp Asn Phe Arg Lys Asn Phe Arg Ser Ile Leu Arg Cys
 325 330

(18) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

ACGAATTCAG CCATGGTCCT TGAGGTGAGT GACCACCAAG TGCTAAAT

48

30 (19) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GAGGATCCTG GAATGCGGGG AAGTCAG

27

(20) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

5

(A) LENGTH: 1107 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

10 ATGGTCCTTG AGGTGAGTGA CCACCAAGTG CTAAATGACG CCGAGGTTGC CGCCCTCCTG 60
GAGAACTTCA GCTCTTCCTA TGACTATGGA GAAAACGAGA GTGACTCGTG CTGTACCTCC 120
CCGCCCTGCC CACAGGACTT CAGCCTGAAC TTCGACCGGG CCTTCCTGCC AGCCCTCTAC 180
AGCCTCCTCT TTCTGCTGGG GCTGCTGGGC AACGGCGCGG TGGCAGCCGT GCTGCTGAGC 240
CGGCGGACAG CCCTGAGCAG CACCGACACC TTCCTGCTCC ACCTAGCTGT AGCAGACACG 300
15 CTGCTGGTGC TGACACTGCC GCTCTGGGCA GTGGACGCTG CCGTCCAGTG GGTCTTTGGC 360
TCTGGCCTCT GCAAAGTGGC AGGTGCCCTC TTCAACATCA ACTTCTACGC AGGAGCCCTC 420
CTGCTGGCCT GCATCAGCTT TGACCGCTAC CTGAACATAG TTCATGCCAC CCAGCTCTAC 480
CGCCGGGGGC CCCC GGCCCG CGTGACCCTC ACCTGCCTGG CTGTCTGGGG GCTCTGCCTG 540
CTTTTCGCCC TCCCAGACTT CATCTTCCTG TCGGCCACAC ACGACGAGCG CCTCAACGCC 600
20 ACCCACTGCC AATACAACTT CCCACAGGTG GGCCGCACGG CTCTGCGGGT GCTGCAGCTG 660
GTGGCTGGCT TTCTGCTGCC CCTGCTGGTC ATGGCCTACT GCTATGCCCA CATCCTGGCC 720
GTGCTGCTGG TTTCCAGGGG CCAGCGGCGC CTGCGGGCCA TGC GGCTGGT GGTGGTGGTC 780
GTGGTGGCCT TTGCCCTCTG CTGGACCCCC TATCACCTGG TGGTGTGGT GGACATCCTC 840
ATGGACCTGG GCGCTTTGGC CCGCAACTGT GGCCGAGAAA GCAGGGTAGA CGTGGCCAAG 900
25 TCGGTACCT CAGGCCTGGG CTACATGCAC TGCTGCCTCA ACCCGCTGCT CTATGCCTTT 960
GTAGGGGTCA AGTTCCGGGA GCGGATGTGG ATGCTGCTCT TGCGCCTGGG CTGCCCCAAC 1020
CAGAGAGGGC TCCAGAGGCA GCCATCGTCT TCCCGCCGGG ATTCATCCTG GTCTGAGACC 1080
TCAGAGGCCT CTTACTCGGG CTTGTGA 1107

(21) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 368 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant.

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Met Val Leu Glu Val Ser Asp His Gln Val Leu Asn Asp Ala Glu Val
 1 5 10 15
 10 Ala Ala Leu Leu Glu Asn Phe Ser Ser Ser Tyr Asp Tyr Gly Glu Asn
 20 25 30
 Glu Ser Asp Ser Cys Cys Thr Ser Pro Pro Cys Pro Gln Asp Phe Ser
 35 40 45
 15 Leu Asn Phe Asp Arg Ala Phe Leu Pro Ala Leu Tyr Ser Leu Leu Phe
 50 55 60
 Leu Leu Gly Leu Leu Gly Asn Gly Ala Val Ala Val Leu Leu Ser
 65 70 75 80
 Arg Arg Thr Ala Leu Ser Ser Thr Asp Thr Phe Leu Leu His Leu Ala
 85 90 95
 20 Val Ala Asp Thr Leu Leu Val Leu Thr Leu Pro Leu Trp Ala Val Asp
 100 105 110
 Ala Ala Val Gln Trp Val Phe Gly Ser Gly Leu Cys Lys Val Ala Gly
 115 120 125
 25 Ala Leu Phe Asn Ile Asn Phe Tyr Ala Gly Ala Leu Leu Leu Ala Cys
 130 135 140
 Ile Ser Phe Asp Arg Tyr Leu Asn Ile Val His Ala Thr Gln Leu Tyr
 145 150 155 160
 Arg Arg Gly Pro Pro Ala Arg Val Thr Leu Thr Cys Leu Ala Val Trp
 165 170 175
 30 Gly Leu Cys Leu Leu Phe Ala Leu Pro Asp Phe Ile Phe Leu Ser Ala
 180 185 190
 His His Asp Glu Arg Leu Asn Ala Thr His Cys Gln Tyr Asn Phe Pro
 195 200 205
 35 Gln Val Gly Arg Thr Ala Leu Arg Val Leu Gln Leu Val Ala Gly Phe
 210 215 220
 Leu Leu Pro Leu Leu Val Met Ala Tyr Cys Tyr Ala His Ile Leu Ala
 225 230 235 240

[illegible]

(22) INFORMATION FOR SEQ ID NO:21:

20

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

25 TTAAGCTTGA CCTAATGCCA TCTTGTGTCC 30

(23) INFORMATION FOR SEQ ID NO:22:

30

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTGGATCCAA AAGAACCATG CACCTCAGAG 30

35 (24) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1074 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
ATGGCTGATG ACTATGGCTC TGAATCCACA TCTTCCATGG AAGACTACGT TAACTTCAAC      60
TTCAC TGACT TCTACTGTGA GAAAAACAAT GTCAGGCAGT TTGCGAGCCA TTCCTCCCA      120
CCCTTG TACT GGCTCGTGT CATCGTGGGT GCCTTGGGCA ACAGTCTTGT TATCCTTGTC      180
10 TACTGG TACT GCACAAGAGT GAAGACCATG ACCGACATGT TCCTTTTGAA TTTGGCAATT      240
GCTGAC CTCC TCTTTCTTGT CACTCTTCCC TTCTGGGCCA TTGCTGCTGC TGACCAGTGG      300
AAGTTC CAGA CCTTCATGTG CAAGGTGGTC AACAGCATGT ACAAGATGAA CTTCTACAGC      360
TGTGTG TTGC TGATCATGTG CATCAGCGTG GACAGGTACA TTGCCATTGC CCAGGCCATG      420
AGAGCA CATA CTTGGAGGGA GAAAAGGCTT TTGTACAGCA AAATGGTTTG CTTTACCATC      480
15 TGGGTAT TGG CAGCTGCTCT CTGCATCCCA GAAATCTTAT ACAGCCAAAT CAAGGAGGAA      540
TCCGGC ATTG CTATCTGCAC CATGGTTTAC CCTAGCGATG AGAGCACCAA ACTGAAGTCA      600
GCTGTCT TGA CCCTGAAGGT CATTCTGGGG TTCTTCCTTC CCTTCGTGGT CATGGCTTGC      660
TGCTATA CCA TCATCATTCA CACCCTGATA CAAGCCAAGA AGTCTTCCAA GCACAAAGCC      720
CTAAAAG TGA CCATCACTGT CCTGACCGTC TTTGTCTTGT CTCAGTTTCC CTACAACTGC      780
20 ATTTTG TTGG TGCAGACCAT TGACGCCTAT GCCATGTTCA TCTCCAAC TGCCGTTTCC      840
ACCAAC ATTG ACATCTGCTT CCAGGTCACC CAGACCATCG CCTTCTTCCA CAGTTGCCTG      900
AACCCT GTTC TCTATGTTTT TGTGGGTGAG AGATTCCGCC GGGATCTCGT GAAAACCCTG      960
AAGAACT TGG GTTGCATCAG CCAGGCCAG TGGGTTTCAT TTACAAGGAG AGAGGGAAGC     1020
TTGAAG CTGT CGTCTATGTT GCTGGAGACA ACCTCAGGAG CACTCTCCCT CTGA      1074
```

25 (25) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 357 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

30

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

	Met	Ala	Asp	Asp	Tyr	Gly	Ser	Glu	Ser	Thr	Ser	Ser	Met	Glu	Asp	Tyr	
	1				5					10					15		
5	Val	Asn	Phe	Asn	Phe	Thr	Asp	Phe	Tyr	Cys	Glu	Lys	Asn	Asn	Val	Arg	
				20					25					30			
	Gln	Phe	Ala	Ser	His	Phe	Leu	Pro	Pro	Leu	Tyr	Trp	Leu	Val	Phe	Ile	
				35				40					45				
	Val	Gly	Ala	Leu	Gly	Asn	Ser	Leu	Val	Ile	Leu	Val	Tyr	Trp	Tyr	Cys	
		50				55						60					
10	Thr	Arg	Val	Lys	Thr	Met	Thr	Asp	Met	Phe	Leu	Leu	Asn	Leu	Ala	Ile	
	65					70					75				80		
	Ala	Asp	Leu	Leu	Phe	Leu	Val	Thr	Leu	Pro	Phe	Trp	Ala	Ile	Ala	Ala	
					85					90				95			
15	Ala	Asp	Gln	Trp	Lys	Phe	Gln	Thr	Phe	Met	Cys	Lys	Val	Val	Asn	Ser	
				100					105					110			
	Met	Tyr	Lys	Met	Asn	Phe	Tyr	Ser	Cys	Val	Leu	Leu	Ile	Met	Cys	Ile	
			115					120					125				
	Ser	Val	Asp	Arg	Tyr	Ile	Ala	Ile	Ala	Gln	Ala	Met	Arg	Ala	His	Thr	
		130				135						140					
20	Trp	Arg	Glu	Lys	Arg	Leu	Leu	Tyr	Ser	Lys	Met	Val	Cys	Phe	Thr	Ile	
	145					150					155					160	
	Trp	Val	Leu	Ala	Ala	Ala	Leu	Cys	Ile	Pro	Glu	Ile	Leu	Tyr	Ser	Gln	
				165						170				175			
25	Ile	Lys	Glu	Glu	Ser	Gly	Ile	Ala	Ile	Cys	Thr	Met	Val	Tyr	Pro	Ser	
				180					185					190			
	Asp	Glu	Ser	Thr	Lys	Leu	Lys	Ser	Ala	Val	Leu	Thr	Leu	Lys	Val	Ile	
			195					200					205				
	Leu	Gly	Phe	Phe	Leu	Pro	Phe	Val	Val	Met	Ala	Cys	Cys	Tyr	Thr	Ile	
		210					215					220					
30	Ile	Ile	His	Thr	Leu	Ile	Gln	Ala	Lys	Lys	Ser	Ser	Lys	His	Lys	Ala	
	225					230					235					240	
	Leu	Lys	Val	Thr	Ile	Thr	Val	Leu	Thr	Val	Phe	Val	Leu	Ser	Gln	Phe	
				245						250				255			
35	Pro	Tyr	Asn	Cys	Ile	Leu	Leu	Val	Gln	Thr	Ile	Asp	Ala	Tyr	Ala	Met	
			260						265					270			
	Phe	Ile	Ser	Asn	Cys	Ala	Val	Ser	Thr	Asn	Ile	Asp	Ile	Cys	Phe	Gln	
			275						280				285				

20

Val Thr Gln Thr Ile Ala Phe Phe His Ser Cys Leu Asn Pro Val Leu
 290 295 300

Tyr Val Phe Val Gly Glu Arg Phe Arg Arg Asp Leu Val Lys Thr Leu
 305 310 315 320

5 Lys Asn Leu Gly Cys Ile Ser Gln Ala Gln Trp Val Ser Phe Thr Arg
 325 330 335

Arg Glu Gly Ser Leu Lys Leu Ser Ser Met Leu Leu Glu Thr Thr Ser
 340 345 350

10 Gly Ala Leu Ser Leu
 355

(26) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1110 base pairs
 (B) TYPE: nucleic acid
 15 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATGGCCTCAT CGACCACTCG GGGCCCCAGG GTTCTGACT TATTTTCTGG GCTGCCGCCG 60

20 GCGGTCACAA CTCCCGCCAA CCAGAGCGCA GAGGCCTCGG CGGGCAACGG GTCGGTGGCT 120

GGCGCGGACG CTCCAGCCGT CACGCCCTTC CAGAGCCTGC AGCTGGTGCA TCAGCTGAAG 180

GGGCTGATCG TGCTGCTCTA CAGCGTCGTG GTGGTCGTGG GGCTGGTGGG CAACTGCCTG 240

CTGGTGCTGG TGATCGCGCG GGTGCCGCGG CTGCACAACG TGACGAACTT CCTCATCGGC 300

AACCTGGCCT TGTCGACGT GCTCATGTGC ACCGCCTGCG TGCCGCTCAC GCTGGCCTAT 360

25 GCCTTCGAGC CACGCGGCTG GGTGTTCTGGC GGCGGCCTGT GCCACCTGGT CTTCTTCCTG 420

CAGCCGGTCA CCGTCTATGT GTCGGTGTTC ACGCTCACCA CCATCGCAGT GGACCGCTAC 480

GTCGTGCTGG TGCACCCGCT GAGGCGCGCA TCTCGCTGCG CCTCAGCCTA CGCTGTGCTG 540

GCCATCTGGG CGCTGTCCGC GTGCTGGCG CTGCCGCCCG CCGTGACAC CTATCACGTG 600

GAGCTCAAGC CGCAGGACGT GCGCCTCTGC GAGGAGTTCT GGGGCTCCCA GGAGCGCCAG 660

30 CGCCAGCTCT ACGCCTGGGG GCTGCTGCTG GTCACCTACC TGCTCCCTCT GCTGGTCATC 720

CTCCTGTCTT ACGTCCGGGT GTCAGTGAAG CTCCGCAACC GCGTGGTGCC GGGCTGCGTG 780

ACCCAGAGCC AGGCCGACTG GGACCGCGCT CGGCGCCGGC GCACCTTCTG CTTGCTGGTG 840

GTGGTCGTGG TGGTGTTCGC CGTCTGCTGG CTGCCGCTGC ACGTCTTCAA CCTGCTGCGG 900
 GACCTCGACC CCCACGCCAT CGACCCTTAC GCCTTTGGGC TGGTGCAGCT GCTCTGCCAC 960
 TGGCTCGCCA TGAGTTCGGC CTGCTACAAC CCCTTCATCT ACCCCTGGCT GCACGACAGC 1020
 TTCCGCGAGG AGCTGCGCAA ACTGTTGGTC GCTTGGCCCC GCAAGATAGC CCCCATGGC 1080
 5 CAGAATATGA CCGTCAGCGT GGTCACTGA 1110

(27) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 369 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

15 Met Ala Ser Ser Thr Thr Arg Gly Pro Arg Val Ser Asp Leu Phe Ser
 1 5 10 15
 Gly Leu Pro Pro Ala Val Thr Thr Pro Ala Asn Gln Ser Ala Glu Ala
 20 25 30
 Ser Ala Gly Asn Gly Ser Val Ala Gly Ala Asp Ala Pro Ala Val Thr
 35 40 45
 20 Pro Phe Gln Ser Leu Gln Leu Val His Gln Leu Lys Gly Leu Ile Val
 50 55 60
 Leu Leu Tyr Ser Val Val Val Val Val Gly Leu Val Gly Asn Cys Leu
 65 70 75 80
 25 Leu Val Leu Val Ile Ala Arg Val Pro Arg Leu His Asn Val Thr Asn
 85 90 95
 Phe Leu Ile Gly Asn Leu Ala Leu Ser Asp Val Leu Met Cys Thr Ala
 100 105 110
 Cys Val Pro Leu Thr Leu Ala Tyr Ala Phe Glu Pro Arg Gly Trp Val
 115 120 125
 30 Phe Gly Gly Gly Leu Cys His Leu Val Phe Phe Leu Gln Pro Val Thr
 130 135 140
 Val Tyr Val Ser Val Phe Thr Leu Thr Thr Ile Ala Val Asp Arg Tyr
 145 150 155 160
 35 Val Val Leu Val His Pro Leu Arg Arg Ala Ser Arg Cys Ala Ser Ala
 165 170 175

22

Tyr Ala Val Leu Ala Ile Trp Ala Leu Ser Ala Val Leu Ala Leu Pro
 180 185 190
 Pro Ala Val His Thr Tyr His Val Glu Leu Lys Pro His Asp Val Arg
 195 200 205
 5 Leu Cys Glu Glu Phe Trp Gly Ser Gln Glu Arg Gln Arg Gln Leu Tyr
 210 215 220
 Ala Trp Gly Leu Leu Leu Val Thr Tyr Leu Leu Pro Leu Leu Val Ile
 225 230 235 240
 10 Leu Leu Ser Tyr Val Arg Val Ser Val Lys Leu Arg Asn Arg Val Val
 245 250 255
 Pro Gly Cys Val Thr Gln Ser Gln Ala Asp Trp Asp Arg Ala Arg Arg
 260 265 270
 Arg Arg Thr Phe Cys Leu Leu Val Val Val Val Val Val Phe Ala Val
 275 280 285
 15 Cys Trp Leu Pro Leu His Val Phe Asn Leu Leu Arg Asp Leu Asp Pro
 290 295 300
 His Ala Ile Asp Pro Tyr Ala Phe Gly Leu Val Gln Leu Leu Cys His
 305 310 315 320
 20 Trp Leu Ala Met Ser Ser Ala Cys Tyr Asn Pro Phe Ile Tyr Ala Trp
 325 330 335
 Leu His Asp Ser Phe Arg Glu Glu Leu Arg Lys Leu Leu Val Ala Trp
 340 345 350
 Pro Arg Lys Ile Ala Pro His Gly Gln Asn Met Thr Val Ser Val Val
 355 360 365
 25 Ile

(28) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1083 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
 35 ATGGACCCAG AAGAACTTC AGTTTATTTG GATTATTACT ATGCTACGAG CCCAACTCT 60
 GACATCAGGG AGACCCACTC CCATGTTCTT TACACCTCTG TCTTCCTTCC AGTCTTTTAC 120

ACAGCTGTGT TCCTGACTGG AGTGCTGGGG AACCTTGTTT TCATGGGAGC GTTGCAATTC 180
 AAACCCGGCA GCCGAAGACT GATCGACATC TTTATCATCA ATCTGGCTGC CTCTGACTTC 240
 ATTTTCTTGG TCACATTGCC TCTCTGGGTG GATAAAGAAG CATCTCTAGG ACTGTGGAGG 300
 ACGGGCTCCT TCCTGTGCAA AGGGAGCTCC TACATGATCT CCGTCAATAT GCACTGCAGT 360
 5 GTCCTCCTGC TCACTTGCAAT GAGTGTGAC CGCTACCTGG CCATTGTGTG GCCAGTCGTA 420
 TCCAGGAAAT TCAGAAGGAC AGACTGTGCA TATGTAGTCT GTGCCAGCAT CTGGTTTATC 480
 TCCTGCCTGC TGGGGTTGCC TACTCTTCTG TCCAGGGAGC TCACGCTGAT TGATGATAAG 540
 CCATACTGTG CAGAGAAAAA GGCAACTCCA ATTAACTCA TATGGTCCCT GGTGGCCTTA 600
 ATTTTCACCT TTTTGTGCC TTTGTTGAGC ATTGTGACCT GCTACTGTTG CATTGCAAGG 660
 10 AAGCTGTGTG CCCATTACCA GCAATCAGGA AAGCACAACA AAAAGCTGAA GAAATCTATA 720
 AAGATCATCT TTATTGTCGT GGCAGCCTTT CTTGTCTCCT GGCTGCCCTT CAATACTTTC 780
 AAGTTCCTGG CCATTGTCTC TGGGTTGCGG CAAGAACACT ATTTACCCTC AGCTATTCTT 840
 CAGCTTGGTA TGGAGGTGAG TGGACCCTTG GCATTGCGA ACAGCTGTGT CAACCCCTTC 900
 ATTTACTATA TCTTCGACAG CTACATCCGC CGGGCCATTG TCCACTGCTT GTGCCCTTGC 960
 15 CTGAAAAACT ATGACTTTGG GAGTAGCACT GAGACATCAG ATAGTCACCT CACTAAGGCT 1020
 CTCTCCACCT TCATTCATGC AGAAGATTTT GCCAGGAGGA GGAAGAGGTC TGTGTCACTC 1080
 TAA 1083

(29) INFORMATION FOR SEQ ID NO:28:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 360 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant
 (ii) MOLECULE TYPE: protein
- 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
 Met Asp Pro Glu Glu Thr Ser Val Tyr Leu Asp Tyr Tyr Tyr Ala Thr
 1 5 10 15
 Ser Pro Asn Ser Asp Ile Arg Glu Thr His Ser His Val Pro Tyr Thr
 20 25 30
 30 Ser Val Phe Leu Pro Val Phe Tyr Thr Ala Val Phe Leu Thr Gly Val
 35 40 45

24

Leu Gly Asn Leu Val Leu Met Gly Ala Leu His Phe Lys Pro Gly Ser
 50 55 60

Arg Arg Leu Ile Asp Ile Phe Ile Ile Asn Leu Ala Ala Ser Asp Phe
 65 70 75 80

5 Ile Phe Leu Val Thr Leu Pro Leu Trp Val Asp Lys Glu Ala Ser Leu
 85 90 95

Gly Leu Trp Arg Thr Gly Ser Phe Leu Cys Lys Gly Ser Ser Tyr Met
 100 105 110

10 Ile Ser Val Asn Met His Cys Ser Val Leu Leu Leu Thr Cys Met Ser
 115 120 125

Val Asp Arg Tyr Leu Ala Ile Val Trp Pro Val Val Ser Arg Lys Phe
 130 135 140

Arg Arg Thr Asp Cys Ala Tyr Val Val Cys Ala Ser Ile Trp Phe Ile
 145 150 155 160

15 Ser Cys Leu Leu Gly Leu Pro Thr Leu Leu Ser Arg Glu Leu Thr Leu
 165 170 175

Ile Asp Asp Lys Pro Tyr Cys Ala Glu Lys Lys Ala Thr Pro Ile Lys
 180 185 190

20 Leu Ile Trp Ser Leu Val Ala Leu Ile Phe Thr Phe Phe Val Pro Leu
 195 200 205

Leu Ser Ile Val Thr Cys Tyr Cys Cys Ile Ala Arg Lys Leu Cys Ala
 210 215 220

His Tyr Gln Gln Ser Gly Lys His Asn Lys Lys Leu Lys Lys Ser Ile
 225 230 235 240

25 Lys Ile Ile Phe Ile Val Val Ala Ala Phe Leu Val Ser Trp Leu Pro
 245 250 255

Phe Asn Thr Phe Lys Phe Leu Ala Ile Val Ser Gly Leu Arg Gln Glu
 260 265 270

30 His Tyr Leu Pro Ser Ala Ile Leu Gln Leu Gly Met Glu Val Ser Gly
 275 280 285

Pro Leu Ala Phe Ala Asn Ser Cys Val Asn Pro Phe Ile Tyr Tyr Ile
 290 295 300

Phe Asp Ser Tyr Ile Arg Arg Ala Ile Val His Cys Leu Cys Pro Cys
 305 310 315 320

35 Leu Lys Asn Tyr Asp Phe Gly Ser Ser Thr Glu Thr Ser Asp Ser His
 325 330 335

Leu Thr Lys Ala Leu Ser Thr Phe Ile His Ala Glu Asp Phe Ala Arg

25

340

345

350

Arg Arg Lys Arg Ser Val Ser Leu
355 360

(30) INFORMATION FOR SEQ ID NO:29:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- 10 (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTAGAATTCT GACTCCAGCC AAAGCATGAA T

31

(31) INFORMATION FOR SEQ ID NO:30:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GCTGGATCCT AAACAGTCTG CGCTCGGCCT

30

(32) INFORMATION FOR SEQ ID NO:31:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1020 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

30 ATGAATGGCC TTGAAGTGGC TCCCCAGGT CTGATCACCA ACTTCTCCCT GGCCACGGCA 60
GAGCAATGTG GCCAGGAGAC GCCACTGGAG AACATGCTGT TCGCCTCCTT CTACCTTCTG 120
GATTTTATCC TGGCTTTAGT TGGCAATACC CTGGCTCTGT GGCTTTTCAT CCGAGACCAC 180
AAGTCCGGGA CCCC GGCCAA CGTGTTCTCTG ATGCATCTGG CCGTGGCCGA CTTGTCTGTC 240
GTGCTGGTCC TGCCACCCG CCTGGTCTAC CACTTCTCTG GGAACCACTG GCCATTGGG 300

GAAATCGCAT GCCGTCTCAC CGGCTTCCTC TTCTACCTCA ACATGTACGC CAGCATCTAC 360
 TTCCTCACCT GCATCAGCGC CGACCGTTTC CTGGCCATTG TGCACCCGGT CAAGTCCCTC 420
 AAGCTCCGCA GGCCCTCTA CGCACACCTG GCCTGTGCCT TCCTGTGGGT GGTGGTGGCT 480
 GTGGCCATGG CCCCCTGCT GGTGAGCCCA CAGACCGTGC AGACCAACCA CACGGTGGTC 540
 5 TGCCTGCAGC TGTACCGGA GAAGGCCTCC CACCATGCCC TGGTGTCCCT GGCAGTGGCC 600
 TTCACCTTCC CGTTCATCAC CACGGTCACC TGCTACCTGC TGATCATCCG CAGCCTGCGG 660
 CAGGGCCTGC GTGTGGAGAA GCGCCTCAAG ACCAAGGCAG TCGCATGAT CGCCATAGTG 720
 CTGGCCATCT TCCTGGTCTG CTTCTGTGCC TACCACGTCA ACCGCTCCGT CTACGTGCTG 780
 CACTACCGCA GCCATGGGGC CTCCTGCGCC ACCCAGCGCA TCCTGGCCCT GGCAAACCGC 840
 10 ATCACCTCCT GCCTCACCAG CCTCAACGGG GCACTCGACC CCATCATGTA TTTCTTCGTG 900
 GCTGAGAAGT TCCGCCACGC CCTGTGCAAC TTGCTCTGTG GCAAAAGGCT CAAGGGCCCCG 960
 CCCCCAGCT TCGAAGGGAA AACCAACGAG AGCTCGCTGA GTGCCAAGTC AGAGCTGTGA 1020

(33) INFORMATION FOR SEQ ID NO:32:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

 (ii) MOLECULE TYPE: protein

 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

 Met Asn Gly Leu Glu Val Ala Pro Pro Gly Leu Ile Thr Asn Phe Ser
 1 5 10 15
 Leu Ala Thr Ala Glu Gln Cys Gly Gln Glu Thr Pro Leu Glu Asn Met
 20 25 30
 25 Leu Phe Ala Ser Phe Tyr Leu Leu Asp Phe Ile Leu Ala Leu Val Gly
 35 40 45
 Asn Thr Leu Ala Leu Trp Leu Phe Ile Arg Asp His Lys Ser Gly Thr
 50 55 60
 30 Pro Ala Asn Val Phe Leu Met His Leu Ala Val Ala Asp Leu Ser Cys
 65 70 75 80
 Val Leu Val Leu Pro Thr Arg Leu Val Tyr His Phe Ser Gly Asn His
 85 90 95
 Trp Pro Phe Gly Glu Ile Ala Cys Arg Leu Thr Gly Phe Leu Phe Tyr

27

	100	105	110
	Leu Asn Met Tyr Ala Ser Ile Tyr Phe Leu Thr Cys Ile Ser Ala Asp		
	115	120	125
5	Arg Phe Leu Ala Ile Val His Pro Val Lys Ser Leu Lys Leu Arg Arg		
	130	135	140
	Pro Leu Tyr Ala His Leu Ala Cys Ala Phe Leu Trp Val Val Val Ala		
	145	150	155
	Val Ala Met Ala Pro Leu Leu Val Ser Pro Gln Thr Val Gln Thr Asn		
	165	170	175
10	His Thr Val Val Cys Leu Gln Leu Tyr Arg Glu Lys Ala Ser His His		
	180	185	190
	Ala Leu Val Ser Leu Ala Val Ala Phe Thr Phe Pro Phe Ile Thr Thr		
	195	200	205
15	Val Thr Cys Tyr Leu Leu Ile Ile Arg Ser Leu Arg Gln Gly Leu Arg		
	210	215	220
	Val Glu Lys Arg Leu Lys Thr Lys Ala Val Arg Met Ile Ala Ile Val		
	225	230	235
	Leu Ala Ile Phe Leu Val Cys Phe Val Pro Tyr His Val Asn Arg Ser		
	245	250	255
20	Val Tyr Val Leu His Tyr Arg Ser His Gly Ala Ser Cys Ala Thr Gln		
	260	265	270
	Arg Ile Leu Ala Leu Ala Asn Arg Ile Thr Ser Cys Leu Thr Ser Leu		
	275	280	285
25	Asn Gly Ala Leu Asp Pro Ile Met Tyr Phe Phe Val Ala Glu Lys Phe		
	290	295	300
	Arg His Ala Leu Cys Asn Leu Leu Cys Gly Lys Arg Leu Lys Gly Pro		
	305	310	315
	Pro Pro Ser Phe Glu Gly Lys Thr Asn Glu Ser Ser Leu Ser Ala Lys		
	325	330	335
30	Ser Glu Leu		

(34) INFORMATION FOR SEQ ID NO:33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

ATAAGATGAT CACCCTGAAC AATCAAGAT

29

(35) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TCCGAATTCA TAACATTTC A CTGTTTATAT TGC

33

(36) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- 15 (A) LENGTH: 996 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

20 ATGATCACCC TGAACAATCA AGATCAACCT GTCACCTTTA ACAGCTCACA TCCAGATGAA 60
TACAAAATTG CAGCCCTTGT CTTCTATAGC TGTATCTTCA TAATTGGATT ATTTGTTAAC 120
ATCACTGCAT TATGGGTTTT CAGTTGTACC ACCAAGAAGA GAACCACGGT AACCATCTAT 180
ATGATGAATG TGGCATTAGT GGACTTGATA TTTATAATGA CTTTACCCTT TCGAATGTTT 240
TATTATGCAA AAGATGCATG GCCATTTGGA GAGTACTTCT GCCAGATTAT TGGAGCTCTC 300
25 ACAGTGTTTT ACCCAAGCAT TGCTTTATGG CTTCTTGCCT TTATTAGTGC TGACAGATAC 360
ATGGCCATTG TACAGCCGAA GTACGCCAAA GAACTTAAAA ACACGTGCAA AGCCGTGCTG 420
GCGTGTGTGG GAGTCTGGAT AATGACCCTG ACCACGACCA CCCCTCTGCT ACTGCTCTAT 480
AAAGACCCAG ATAAAGACTC CACTCCCGCC ACCTGCCTCA AGATTTCTGA CATCATCTAT 540
CTAAAAGCTG TGAACGTGCT GAACCTCACT CGACTGACAT TTTTTTCTT GATTCTTTG 600
30 TTCATCATGA TTGGGTGCTA CTTGGTCATT ATTCATAATC TCCTTCACGG CAGGACGTCT 660
AAGCTGAAAC CCAAAGTCAA GGAGAAGTCC ATAAGGATCA TCATCACGCT GCTGGTGCAG 720

GTGCTCGTCT GCTTTATGCC CTTCCACATC TGTTTCGCTT TCCTGATGCT GGGAACGGGG 780
 GAGAACAGTT ACAATCCCTG GGGAGCCTTT ACCACCTTCC TCATGAACCT CAGCACGTGT 840
 CTGGATGTGA TTCTCTACTA CATCGTTTCA AAACAATTTC AGGCTCGAGT CATTAGTGTC 900
 ATGCTATACC GTAATTACCT TCGAAGCCTG CGCAGAAAAA GTTTCGATC TGGTAGTCTA 960
 5 AGGTCATAA GCAATATAAA CAGTGAAATG TTATGA 996

(37) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 331 amino acids
 (B) TYPE: amino acid
 10 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

15 Met Ile Thr Leu Asn Asn Gln Asp Gln Pro Val Thr Phe Asn Ser Ser
 1 5 10 15
 His Pro Asp Glu Tyr Lys Ile Ala Ala Leu Val Phe Tyr Ser Cys Ile
 20 25 30
 Phe Ile Ile Gly Leu Phe Val Asn Ile Thr Ala Leu Trp Val Phe Ser
 35 40 45
 20 Cys Thr Thr Lys Lys Arg Thr Thr Val Thr Ile Tyr Met Met Asn Val
 50 55 60
 Ala Leu Val Asp Leu Ile Phe Ile Met Thr Leu Pro Phe Arg Met Phe
 65 70 75 80
 25 Tyr Tyr Ala Lys Asp Ala Trp Pro Phe Gly Glu Tyr Phe Cys Gln Ile
 85 90 95
 Ile Gly Ala Leu Thr Val Phe Tyr Pro Ser Ile Ala Leu Trp Leu Leu
 100 105 110
 Ala Phe Ile Ser Ala Asp Arg Tyr Met Ala Ile Val Gln Pro Lys Tyr
 115 120 125
 30 Ala Lys Glu Leu Lys Asn Thr Cys Lys Ala Val Leu Ala Cys Val Gly
 130 135 140
 Val Trp Ile Met Thr Leu Thr Thr Thr Thr Pro Leu Leu Leu Tyr
 145 150 155 160
 35 Lys Asp Pro Asp Lys Asp Ser Thr Pro Ala Thr Cys Leu Lys Ile Ser
 165 170 175

30

Asp Ile Ile Tyr Leu Lys Ala Val Asn Val Leu Asn Leu Thr Arg Leu
 180 185 190
 Thr Phe Phe Phe Leu Ile Pro Leu Phe Ile Met Ile Gly Cys Tyr Leu
 195 200 205
 5 Val Ile Ile His Asn Leu Leu His Gly Arg Thr Ser Lys Leu Lys Pro
 210 215 220
 Lys Val Lys Glu Lys Ser Ile Arg Ile Ile Ile Thr Leu Leu Val Gln
 225 230 235 240
 10 Val Leu Val Cys Phe Met Pro Phe His Ile Cys Phe Ala Phe Leu Met
 245 250 255
 Leu Gly Thr Gly Glu Asn Ser Tyr Asn Pro Trp Gly Ala Phe Thr Thr
 260 265 270
 Phe Leu Met Asn Leu Ser Thr Cys Leu Asp Val Ile Leu Tyr Tyr Ile
 275 280 285
 15 Val Ser Lys Gln Phe Gln Ala Arg Val Ile Ser Val Met Leu Tyr Arg
 290 295 300
 Asn Tyr Leu Arg Ser Leu Arg Arg Lys Ser Phe Arg Ser Gly Ser Leu
 305 310 315 320
 20 Arg Ser Leu Ser Asn Ile Asn Ser Glu Met Leu
 325 330

(38) INFORMATION FOR SEQ ID NO:37:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CCAAGCTTCC AGGCCTGGGG TGTGCTGG

28

30 (39) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 35 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

ATGGATCCTG ACCTTCGGCC CCTGGCAGA

29

(40) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 1077 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

10 ATGCCCTCTG TGTCTCCAGC GGGGCCCTCG GCCGGGGCAG TCCCAATGC CACCGCAGTG 60
ACAACAGTGC GGACCAATGC CAGCGGGCTG GAGGTGCCCC TGTTCCACCT GTTTGCCCCG 120
CTGGACGAGG AGCTGCATGG CACCTTCCCA GGCCTGTGCG TGGCGCTGAT GGCGGTGCAC 180
GGAGCCATCT TCCTGGCAGG GCTGGTGCTC AACGGGCTGG CGCTGTACGT CTTCTGCTGC 240
CGCACC CGG CCAAGACACC CTCAGTCATC TACACCATCA ACCTGGTGGT GACCGATCTA 300
15 CTGGTAGGGC TGTCCCTGCC CACGCGCTTC GCTGTGTACT ACGGCGCCAG GGGCTGCCTG 360
CGCTGTGCCT TCCCGCACGT CCTCGGTTAC TTCCTCAACA TGCACTGCTC CATCCTCTTC 420
CTCACCTGCA TCTGCGTGGA CCGCTACCTG GCCATCGTGC GGCCGAAGG CTCCCGCCGC 480
TGCCGCCAGC CTGCCTGTGC CAGGGCCGTG TCGCCTTCG TGTGGCTGGC CGCCGGTGCC 540
GTCACCCTGT CGGTGCTGGG CGTGACAGGC AGCCGGCCCT GCTGCCGTGT CTTTGCGCTG 600
20 ACTGTCCTGG AGTTCCTGCT GCCCCTGCTG GTCATCAGCG TGTTTACCGG CCGCATCATG 660
TGTGCACTGT CGCGGCCGGG TCTGCTCCAC CAGGGTCGCC AGCGCCGCGT GCGGGCCATG 720
CAGCTCCTGC TCACGGTGCT CATCATCTTT CTCGTCTGCT TCACGCCCTT CCACGCCCCG 780
CAAGTGGCCG TGGCGCTGTG GCCCACATG CCACACCACA CGAGCCTCGT GGTCTACCAC 840
GTGGCCGTGA CCCTCAGCAG CCTCAACAGC TGCATGGACC CCATCGTCTA CTGCTTCGTC 900
25 ACCAGTGGCT TCCAGGCCAC CGTCCGAGGC CTCTTCGGCC AGCACGGAGA GCGTGAGCCC 960
AGCAGCGGTG ACGTGGTCAG CATGCACAGG AGCTCCAAGG GCTCAGGCCG TCATCACATC 1020
CTCAGTGCCG GCCCTCACGC CCTCACCAG GCCCTGGCTA ATGGGCCCGA GGCTTAG 1077

(41) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

30

- (A) LENGTH: 358 amino acids

32

(B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met	Pro	Ser	Val	Ser	Pro	Ala	Gly	Pro	Ser	Ala	Gly	Ala	Val	Pro	Asn	
1				5					10					15		
Ala	Thr	Ala	Val	Thr	Thr	Val	Arg	Thr	Asn	Ala	Ser	Gly	Leu	Glu	Val	
			20					25					30			
Pro	Leu	Phe	His	Leu	Phe	Ala	Arg	Leu	Asp	Glu	Glu	Leu	His	Gly	Thr	
		35					40					45				
Phe	Pro	Gly	Leu	Cys	Val	Ala	Leu	Met	Ala	Val	His	Gly	Ala	Ile	Phe	
		50				55					60					
Leu	Ala	Gly	Leu	Val	Leu	Asn	Gly	Leu	Ala	Leu	Tyr	Val	Phe	Cys	Cys	
65				70				75						80		
Arg	Thr	Arg	Ala	Lys	Thr	Pro	Ser	Val	Ile	Tyr	Thr	Ile	Asn	Leu	Val	
				85				90					95			
Val	Thr	Asp	Leu	Leu	Val	Gly	Leu	Ser	Leu	Pro	Thr	Arg	Phe	Ala	Val	
			100					105					110			
Tyr	Tyr	Gly	Ala	Arg	Gly	Cys	Leu	Arg	Cys	Ala	Phe	Pro	His	Val	Leu	
		115					120					125				
Gly	Tyr	Phe	Leu	Asn	Met	His	Cys	Ser	Ile	Leu	Phe	Leu	Thr	Cys	Ile	
		130				135					140					
Cys	Val	Asp	Arg	Tyr	Leu	Ala	Ile	Val	Arg	Pro	Glu	Ala	Pro	Ala	Ala	
145					150					155				160		
Cys	Arg	Gln	Pro	Ala	Cys	Ala	Arg	Ala	Val	Cys	Ala	Phe	Val	Trp	Leu	
				165					170					175		
Ala	Ala	Gly	Ala	Val	Thr	Leu	Ser	Val	Leu	Gly	Val	Thr	Gly	Ser	Arg	
			180					185					190			
Pro	Cys	Cys	Arg	Val	Phe	Ala	Leu	Thr	Val	Leu	Glu	Phe	Leu	Leu	Pro	
		195					200					205				
Leu	Leu	Val	Ile	Ser	Val	Phe	Thr	Gly	Arg	Ile	Met	Cys	Ala	Leu	Ser	
		210				215					220					
Arg	Pro	Gly	Leu	Leu	His	Gln	Gly	Arg	Gln	Arg	Arg	Val	Arg	Ala	Met	
225					230					235				240		
Gln	Leu	Leu	Leu	Thr	Val	Leu	Ile	Ile	Phe	Leu	Val	Cys	Phe	Thr	Pro	
				245					250					255		

33

Phe His Ala Arg Gln Val Ala Val Ala Leu Trp Pro Asp Met Pro His
 260 265 270
 His Thr Ser Leu Val Val Tyr His Val Ala Val Thr Leu Ser Ser Leu
 275 280 285
 5 Asn Ser Cys Met Asp Pro Ile Val Tyr Cys Phe Val Thr Ser Gly Phe
 290 295 300
 Gln Ala Thr Val Arg Gly Leu Phe Gly Gln His Gly Glu Arg Glu Pro
 305 310 315 320
 10 Ser Ser Gly Asp Val Val Ser Met His Arg Ser Ser Lys Gly Ser Gly
 325 330 335
 Arg His His Ile Leu Ser Ala Gly Pro His Ala Leu Thr Gln Ala Leu
 340 345 350
 Ala Asn Gly Pro Glu Ala
 355

15 (42) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GAGAATTCAC TCCTGAGCTC AAGATGAACT

30

(43) INFORMATION FOR SEQ ID NO:42:

25 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

CGGGATCCCC GTAAGTGAGC CACTTCAGAT

30

(44) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 1050 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

ATGAACTCCA CCTTGGATGG TAATCAGAGC AGCCACCCTT TTTGCCTCTT GGCATTGTCG 60
5 TATTTGGAAG CTGTCAATTT TTGCCTTTTG GAAGTATTGA TTATTGTCTT TCTAACTGTA 120
TTGATTATTT CTGGCAACAT CATTGTGATT TTTGTATTTT ACTGTGCACC TTTGTTGAAC 180
CATCACACTA CAAGTTATTT TATCCAGACT ATGGCATATG CTGACCTTTT TGTTGGGGTG 240
AGCTGCGTGG TCCCTTCTTT ATCACTCCTC CATCACCCCC TTCCAGTAGA GGAGTCCTTG 300
ACTTGCCAGA TATTTGGTTT TGTAATATCA GTTCTGAAGA GCGTCTCCAT GGCTTCTCTG 360
10 GCCTGTATCA GCATTGATAG ATACATTGCC ATTACTAAAC CTTTAACCTA TAATACTCTG 420
GTTACACCCT GGAGACTACG CCTGTGTATT TTCCTGATTT GGCTATACTC GACCCTGGTC 480
TTCCTGCCTT CCTTTTTCCTA CTGGGGCAAA CCTGGATATC ATGGAGATGT GTTTCAGTGG 540
TGTGCGGAGT CCTGGCACAC CGACTCCTAC TTCACCCTGT TCATCGTGAT GATGTTATAT 600
GCCCCAGCAG CCCTTATTGT CTGCTTCACC TATTTCAACA TCTTCCGCAT CTGCCAACAG 660
15 CACACAAAGG ATATCAGCGA AAGGCAAGCC CGCTTCAGCA GCCAGAGTGG GGAGACTGGG 720
GAAGTGCAGG CCTGTCCTGA TAAGCGCTAT GCCATGGTCC TGTTTCGAAT CACTAGTGTA 780
TTTTACATCC TCTGGTTGCC ATATATCATC TACTTCTTGT TGGAAAGCTC CACTGGCCAC 840
AGCAACCGCT TCGCATCCTT CTTGACCACC TGGCTTGCTA TTAGTAACAG TTTCTGCAAC 900
TGTGTAATTT ATAGTCTCTC CAACAGTGTA TTCCAAAGAG GACTAAAGCG CCTCTCAGGG 960
20 GCTATGTGTA CTTCTTGTGC AAGTCAGACT ACAGCCAACG ACCCTTACAC AGTTAGAAGC 1020
AAAGGCCCTC TTAATGGATG TCATATCTGA 1050

(45) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 349 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

30 Met Asn Ser Thr Leu Asp Gly Asn Gln Ser Ser His Pro Phe Cys Leu

35

	1		5		10		15									
	Leu	Ala	Phe	Gly	Tyr	Leu	Glu	Thr	Val	Asn	Phe	Cys	Leu	Leu	Glu	Val
				20					25					30		
5	Leu	Ile	Ile	Val	Phe	Leu	Thr	Val	Leu	Ile	Ile	Ser	Gly	Asn	Ile	Ile
			35					40					45			
	Val	Ile	Phe	Val	Phe	His	Cys	Ala	Pro	Leu	Leu	Asn	His	His	Thr	Thr
	50						55					60				
	Ser	Tyr	Phe	Ile	Gln	Thr	Met	Ala	Tyr	Ala	Asp	Leu	Phe	Val	Gly	Val
	65					70					75				80	
10	Ser	Cys	Val	Val	Pro	Ser	Leu	Ser	Leu	Leu	His	His	Pro	Leu	Pro	Val
					85					90					95	
	Glu	Glu	Ser	Leu	Thr	Cys	Gln	Ile	Phe	Gly	Phe	Val	Val	Ser	Val	Leu
				100					105					110		
15	Lys	Ser	Val	Ser	Met	Ala	Ser	Leu	Ala	Cys	Ile	Ser	Ile	Asp	Arg	Tyr
			115					120					125			
	Ile	Ala	Ile	Thr	Lys	Pro	Leu	Thr	Tyr	Asn	Thr	Leu	Val	Thr	Pro	Trp
	130						135					140				
	Arg	Leu	Arg	Leu	Cys	Ile	Phe	Leu	Ile	Trp	Leu	Tyr	Ser	Thr	Leu	Val
	145					150					155					160
20	Phe	Leu	Pro	Ser	Phe	Phe	His	Trp	Gly	Lys	Pro	Gly	Tyr	His	Gly	Asp
					165					170					175	
	Val	Phe	Gln	Trp	Cys	Ala	Glu	Ser	Trp	His	Thr	Asp	Ser	Tyr	Phe	Thr
				180					185					190		
25	Leu	Phe	Ile	Val	Met	Met	Leu	Tyr	Ala	Pro	Ala	Ala	Leu	Ile	Val	Cys
			195					200						205		
	Phe	Thr	Tyr	Phe	Asn	Ile	Phe	Arg	Ile	Cys	Gln	Gln	His	Thr	Lys	Asp
	210						215					220				
	Ile	Ser	Glu	Arg	Gln	Ala	Arg	Phe	Ser	Ser	Gln	Ser	Gly	Glu	Thr	Gly
	225					230					235				240	
30	Glu	Val	Gln	Ala	Cys	Pro	Asp	Lys	Arg	Tyr	Ala	Met	Val	Leu	Phe	Arg
					245					250					255	
	Ile	Thr	Ser	Val	Phe	Tyr	Ile	Leu	Trp	Leu	Pro	Tyr	Ile	Ile	Tyr	Phe
				260					265					270		
35	Leu	Leu	Glu	Ser	Ser	Thr	Gly	His	Ser	Asn	Arg	Phe	Ala	Ser	Phe	Leu
			275					280						285		
	Thr	Thr	Trp	Leu	Ala	Ile	Ser	Asn	Ser	Phe	Cys	Asn	Cys	Val	Ile	Tyr
	290						295					300				

36

Ser Leu Ser Asn Ser Val Phe Gln Arg Gly Leu Lys Arg Leu Ser Gly
 305 310 315 320

Ala Met Cys Thr Ser Cys Ala Ser Gln Thr Thr Ala Asn Asp Pro Tyr
 325 330 335

5 Thr Val Arg Ser Lys Gly Pro Leu Asn Gly Cys His Ile
 340 345

(46) INFORMATION FOR SEQ ID NO:45:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

15 TCCCCCGGGA AAAAAACCAA CTGCTCCAAA 30

(47) INFORMATION FOR SEQ ID NO:46:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TAGGATCCAT TTGAATGTGG ATTTGGTGAA A 31

25 (48) INFORMATION FOR SEQ ID NO:47:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1302 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

ATGTGTTTTT CTCCCATCTT GGAAATCAAC ATGCAGTCTG AATCTAACAT TACAGTGCGA 60

GATGACATTG ATGACATCAA CACCAATATG TACCAACCAC TATCATATCC GTTAAGCTTT 120

35 CAAGTGTCTC TCACCGGATT TCTTATGTGA GAAATTGTGT TGGGACTTGG CAGCAACCTC 180

37

ACTGTATTGG TACTTTACTG CATGAAATCC AACTTAATCA ACTCTGTCAG TAACATTATT 240
 ACAATGAATC TTCATGTACT TGATGTAATA ATTTGTGTGG GATGTATTCC TCTAACTATA 300
 GTTATCCTTC TGCTTTCACT GGAGAGTAAC ACTGCTCTCA TTTGCTGTTT CCATGAGGCT 360
 TGTGTATCTT TTGCAAGTGT CTCAACAGCA ATCAACGTTT TTGCTATCAC TTTGGACAGA 420
 5 TATGACATCT CTGTAAAACC TGCAAACCGA ATTCTGACAA TGGGCAGAGC TGTAATGTTA 480
 ATGATATCCA TTTGGATTTT TTCTTTTTTC TCTTTCCTGA TTCCTTTTAT TGAGGTAAAT 540
 TTTTTCAGTC TTCAAAGTGG AAATACCTGG GAAAACAAGA CACTTTTATG TGTCAGTACA 600
 AATGAATACT ACACTGAACT GGAATGTAT TATCACCTGT TAGTACAGAT CCCAATATTC 660
 TTTTTCAGTC TTGTAGTAAT GTTAATCACA TACACCAAAA TACTTCAGGC TCTTAATATT 720
 10 CGAATAGGCA CAAGATTTTC AACAGGGCAG AAGAAGAAAG CAAGAAAGAA AAAGACAATT 780
 TCTCTAACCA CACAACATGA GGCTACAGAC ATGTCACAAA GCAGTGGTGG GAGAAATGTA 840
 GTCTTTGGTG TAAGAACTTC AGTTTCTGTA ATAATTGCCC TCCGGCGAGC TGTGAAACGA 900
 CACCGTGAAC GACGAGAAAG ACAAAGAGA GTCTTCAGGA TGTCTTTATT GATTATTTCT 960
 ACATTTCTTC TCTGCTGGAC ACCAATTTCT GTTTTAAATA CCACCATTTT ATGTTTAGGC 1020
 15 CCAAGTGACC TTTTAGTAAA ATTAAGATTG TGTTTTTTAG TCATGGCTTA TGGAACAACT 1080
 ATATTTCACT CTCTATTATA TGCAATCACT AGACAAAAAT TTCAAAGGT CTTGAAAAGT 1140
 AAAATGAAAA AGCGAGTTGT TTCTATAGTA GAAGCTGATC CCCTGCCTAA TAATGCTGTA 1200
 ATACACAACT CTTGGATAGA TCCCAAAAGA AACAAAAAAA TTACCTTTGA AGATAGTGAA 1260
 ATAAGAGAAA AACGTTTAGT GCCTCAGGTT GTCACAGACT AG 1302

20 (49) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 433 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - 25 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Met Cys Phe Ser Pro Ile Leu Glu Ile Asn Met Gln Ser Glu Ser Asn
 1 5 10 15

30 Ile Thr Val Arg Asp Asp Ile Asp Asp Ile Asn Thr Asn Met Tyr Gln
 20 25 30

38

Pro Leu Ser Tyr Pro Leu Ser Phe Gln Val Ser Leu Thr Gly Phe Leu
 35 40 45
 Met Leu Glu Ile Val Leu Gly Leu Gly Ser Asn Leu Thr Val Leu Val
 50 55 60
 5 Leu Tyr Cys Met Lys Ser Asn Leu Ile Asn Ser Val Ser Asn Ile Ile
 65 70 75 80
 Thr Met Asn Leu His Val Leu Asp Val Ile Ile Cys Val Gly Cys Ile
 85 90 95
 10 Pro Leu Thr Ile Val Ile Leu Leu Leu Ser Leu Glu Ser Asn Thr Ala
 100 105 110
 Leu Ile Cys Cys Phe His Glu Ala Cys Val Ser Phe Ala Ser Val Ser
 115 120 125
 Thr Ala Ile Asn Val Phe Ala Ile Thr Leu Asp Arg Tyr Asp Ile Ser
 130 135 140
 15 Val Lys Pro Ala Asn Arg Ile Leu Thr Met Gly Arg Ala Val Met Leu
 145 150 155 160
 Met Ile Ser Ile Trp Ile Phe Ser Phe Phe Ser Phe Leu Ile Pro Phe
 165 170 175
 20 Ile Glu Val Asn Phe Phe Ser Leu Gln Ser Gly Asn Thr Trp Glu Asn
 180 185 190
 Lys Thr Leu Leu Cys Val Ser Thr Asn Glu Tyr Tyr Thr Glu Leu Gly
 195 200 205
 Met Tyr Tyr His Leu Leu Val Gln Ile Pro Ile Phe Phe Phe Thr Val
 210 215 220
 25 Val Val Met Leu Ile Thr Tyr Thr Lys Ile Leu Gln Ala Leu Asn Ile
 225 230 235 240
 Arg Ile Gly Thr Arg Phe Ser Thr Gly Gln Lys Lys Lys Ala Arg Lys
 245 250 255
 30 Lys Lys Thr Ile Ser Leu Thr Thr Gln His Glu Ala Thr Asp Met Ser
 260 265 270
 Gln Ser Ser Gly Gly Arg Asn Val Val Phe Gly Val Arg Thr Ser Val
 275 280 285
 Ser Val Ile Ile Ala Leu Arg Arg Ala Val Lys Arg His Arg Glu Arg
 290 295 300
 35 Arg Glu Arg Gln Lys Arg Val Phe Arg Met Ser Leu Leu Ile Ile Ser
 305 310 315 320
 Thr Phe Leu Leu Cys Trp Thr Pro Ile Ser Val Leu Asn Thr Thr Ile

					325					330					335	
	Leu	Cys	Leu	Gly	Pro	Ser	Asp	Leu	Leu	Val	Lys	Leu	Arg	Leu	Cys	Phe
				340					345					350		
5	Leu	Val	Met	Ala	Tyr	Gly	Thr	Thr	Ile	Phe	His	Pro	Leu	Leu	Tyr	Ala
			355					360					365			
	Phe	Thr	Arg	Gln	Lys	Phe	Gln	Lys	Val	Leu	Lys	Ser	Lys	Met	Lys	Lys
		370					375					380				
	Arg	Val	Val	Ser	Ile	Val	Glu	Ala	Asp	Pro	Leu	Pro	Asn	Asn	Ala	Val
	385					390					395					400
10	Ile	His	Asn	Ser	Trp	Ile	Asp	Pro	Lys	Arg	Asn	Lys	Lys	Ile	Thr	Phe
					405					410					415	
	Glu	Asp	Ser	Glu	Ile	Arg	Glu	Lys	Arg	Leu	Val	Pro	Gln	Val	Val	Thr
				420					425					430		

Asp

15

(50) INFORMATION FOR SEQ ID NO:49:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GTGAAGCTTG CCTCTGGTGC CTGCAGGAGG

30

25 (51) INFORMATION FOR SEQ ID NO:50:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GCAGAATTCC CGGTGGCGTG TTGTGGTGCC C

31

(52) INFORMATION FOR SEQ ID NO:51:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1209 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

	ATGTTGTGTC	CTTCCAAGAC	AGATGGCTCA	GGGCACTCTG	GTAGGATTCA	CCAGGAAACT	60
	CATGGAGAAG	GGAAAAGGGA	CAAGATTAGC	AACAGTGAAG	GGAGGGAGAA	TGGTGGGAGA	120
	GGATTCCAGA	TGAACGGTGG	GTCGCTGGAG	GCTGAGCATG	CCAGCAGGAT	GTCAGTTCCTC	180
	AGAGCAAAGC	CCATGTCAAA	CAGCCAACGC	TTGCTCCTTC	TGTCCCCAGG	ATCACCTCCT	240
10	CGCACGGGGA	GCATCTCCTA	CATCAACATC	ATCATGCCTT	CGGTGTTTCG	CACCATCTGC	300
	CTCCTGGGCA	TCATCGGGAA	CTCCACGGTC	ATCTTCGCGG	TCGTGAAGAA	GTCCAAGCTG	360
	CACTGGTGCA	ACAACGTCCC	CGACATCTTC	ATCATCAACC	TCTCGGTAGT	AGATCTCCTC	420
	TTTCTCCTGG	GCATGCCCTT	CATGATCCAC	CAGCTCATGG	GCAATGGGGT	GTGGCACTTT	480
	GGGGAGACCA	TGTGCACCCT	CATCACGGCC	ATGGATGCCA	ATAGTCAGTT	CACCAGCACC	540
15	TACATCCTGA	CCGCCATGGC	CATTGACCGC	TACCTGGCCA	CTGTCCACCC	CATCTCTTCC	600
	ACGAAGTTCC	GGAAGCCCTC	TGTGGCCACC	CTGGTGATCT	GCCTCCTGTG	GGCCCTCTCC	660
	TTCATCAGCA	TCACCCCTGT	GTGGCTGTAT	GCCAGACTCA	TCCCCTTCCC	AGGAGGTGCA	720
	GTGGGCTGCG	GCATACGCCT	GCCCAACCCA	GACACTGACC	TCTACTGGTT	CACCCTGTAC	780
	CAGTTTTTCC	TGGCCTTTGC	CCTGCCTTTT	GTGGTCATCA	CAGCCGCATA	CGTGAGGATC	840
20	CTGCAGCGCA	TGACGTCCTC	AGTGGCCCCC	GCCTCCCAGC	GCAGCATCCG	GCTGCGGACA	900
	AAGAGGGTGA	CCCGCACAGC	CATCGCCATC	TGTCTGGTCT	TCTTTGTGTG	CTGGGCACCC	960
	TACTATGTGC	TACAGCTGAC	CCAGTTGTCC	ATCAGCCGCC	CGACCCTCAC	CTTTGTCTAC	1020
	TTATACAATG	CGGCCATCAG	CTTGGGCTAT	GCCAACAGCT	GCCTCAACCC	CTTTGTGTAC	1080
	ATCGTGCTCT	GTGAGACGTT	CCGCAAACGC	TTGGTCCTGT	CGGTGAAGCC	TGCAGCCCAG	1140
25	GGGCAGCTTC	GCGCTGTCAG	CAACGCTCAG	ACGGCTGACG	AGGAGAGGAC	AGAAAGCAAA	1200
	GGCACCTGA						1209

(53) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 402 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

5 Met Leu Cys Pro Ser Lys Thr Asp Gly Ser Gly His Ser Gly Arg Ile
 1 5 10 15
 His Gln Glu Thr His Gly Glu Gly Lys Arg Asp Lys Ile Ser Asn Ser
 20 25 30
 10 Glu Gly Arg Glu Asn Gly Gly Arg Gly Phe Gln Met Asn Gly Gly Ser
 35 40 45
 Leu Glu Ala Glu His Ala Ser Arg Met Ser Val Leu Arg Ala Lys Pro
 50 55 60
 Met Ser Asn Ser Gln Arg Leu Leu Leu Leu Ser Pro Gly Ser Pro Pro
 65 70 75 80
 15 Arg Thr Gly Ser Ile Ser Tyr Ile Asn Ile Ile Met Pro Ser Val Phe
 85 90 95
 Gly Thr Ile Cys Leu Leu Gly Ile Ile Gly Asn Ser Thr Val Ile Phe
 100 105 110
 20 Ala Val Val Lys Lys Ser Lys Leu His Trp Cys Asn Asn Val Pro Asp
 115 120 125
 Ile Phe Ile Ile Asn Leu Ser Val Val Asp Leu Leu Phe Leu Leu Gly
 130 135 140
 Met Pro Phe Met Ile His Gln Leu Met Gly Asn Gly Val Trp His Phe
 145 150 155 160
 25 Gly Glu Thr Met Cys Thr Leu Ile Thr Ala Met Asp Ala Asn Ser Gln
 165 170 175
 Phe Thr Ser Thr Tyr Ile Leu Thr Ala Met Ala Ile Asp Arg Tyr Leu
 180 185 190
 30 Ala Thr Val His Pro Ile Ser Ser Thr Lys Phe Arg Lys Pro Ser Val
 195 200 205
 Ala Thr Leu Val Ile Cys Leu Leu Trp Ala Leu Ser Phe Ile Ser Ile
 210 215 220
 Thr Pro Val Trp Leu Tyr Ala Arg Leu Ile Pro Phe Pro Gly Gly Ala
 225 230 235 240
 35 Val Gly Cys Gly Ile Arg Leu Pro Asn Pro Asp Thr Asp Leu Tyr Trp
 245 250 255

42

Phe Thr Leu Tyr Gln Phe Phe Leu Ala Phe Ala Leu Pro Phe Val Val
 260 265 270
 Ile Thr Ala Ala Tyr Val Arg Ile Leu Gln Arg Met Thr Ser Ser Val
 275 280 285
 5 Ala Pro Ala Ser Gln Arg Ser Ile Arg Leu Arg Thr Lys Arg Val Thr
 290 295 300
 Arg Thr Ala Ile Ala Ile Cys Leu Val Phe Phe Val Cys Trp Ala Pro
 305 310 315 320
 10 Tyr Tyr Val Leu Gln Leu Thr Gln Leu Ser Ile Ser Arg Pro Thr Leu
 325 330 335
 Thr Phe Val Tyr Leu Tyr Asn Ala Ala Ile Ser Leu Gly Tyr Ala Asn
 340 345 350
 Ser Cys Leu Asn Pro Phe Val Tyr Ile Val Leu Cys Glu Thr Phe Arg
 355 360 365
 15 Lys Arg Leu Val Leu Ser Val Lys Pro Ala Ala Gln Gly Gln Leu Arg
 370 375 380
 Ala Val Ser Asn Ala Gln Thr Ala Asp Glu Glu Arg Thr Glu Ser Lys
 385 390 395 400
 Gly Thr

20

(54) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGCGGATCCA TGGATGTGAC TTCCCAA

27

30

(55) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GGCGGATCCC TACACGGCAC TGCTGAA

27

(56) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1128 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

10 ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAC 60
GCTGCGGGCCC CCAACACCAC CTCCCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC 120
GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
CTCTCGTGCC TCTACACCAT CTTCTCTTC CCCATCGGCT TTGTGGGCAA CATCCTGATC 240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTTCATCAAC 300
15 CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC 360
GAGCGGTACT ACGACATCGC CGTCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC 420
ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC 480
AGGGCCATGC GCTGCAGCCT GTTCCGCACC AAGCACCACG CCCGGCTGAG CTGTGGCCTC 540
ATCTGGATGG CATCCGTGTC AGCCAGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC 600
20 ACCGACGAGG CCTGCTTCTG TTTCCGGAT GTCCGGGAGG TGCAGTGGCT CGAGGTCACG 660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG 720
CTGGTCAGGG CGCACCGGCA CCGTGGGCTG CGGCCCCGGC GGCAGAAGGC GCTCCGCATG 780
ATCTCGCGG TGGTGCTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
25 GCCCACCCCC TCACGGGCCA CATTGTCAAC CTCGCCGCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG 1020
CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTCACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTGA 1128

(57) INFORMATION FOR SEQ ID NO:56:

30

(i) SEQUENCE CHARACTERISTICS:

44

- (A) LENGTH: 375 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

5 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

	Met	Asp	Val	Thr	Ser	Gln	Ala	Arg	Gly	Val	Gly	Leu	Glu	Met	Tyr	Pro	
	1				5					10					15		
10	Gly	Thr	Ala	His	Ala	Ala	Ala	Pro	Asn	Thr	Thr	Ser	Pro	Glu	Leu	Asn	
				20					25					30			
	Leu	Ser	His	Pro	Leu	Leu	Gly	Thr	Ala	Leu	Ala	Asn	Gly	Thr	Gly	Glu	
			35				40						45				
	Leu	Ser	Glu	His	Gln	Gln	Tyr	Val	Ile	Gly	Leu	Phe	Leu	Ser	Cys	Leu	
			50				55					60					
15	Tyr	Thr	Ile	Phe	Leu	Phe	Pro	Ile	Gly	Phe	Val	Gly	Asn	Ile	Leu	Ile	
	65					70					75					80	
	Leu	Val	Val	Asn	Ile	Ser	Phe	Arg	Glu	Lys	Met	Thr	Ile	Pro	Asp	Leu	
				85						90					95		
20	Tyr	Phe	Ile	Asn	Leu	Ala	Val	Ala	Asp	Leu	Ile	Leu	Val	Ala	Asp	Ser	
				100					105					110			
	Leu	Ile	Glu	Val	Phe	Asn	Leu	His	Glu	Arg	Tyr	Tyr	Asp	Ile	Ala	Val	
			115				120						125				
	Leu	Cys	Thr	Phe	Met	Ser	Leu	Phe	Leu	Gln	Val	Asn	Met	Tyr	Ser	Ser	
			130				135						140				
25	Val	Phe	Phe	Leu	Thr	Trp	Met	Ser	Phe	Asp	Arg	Tyr	Ile	Ala	Leu	Ala	
	145					150					155					160	
	Arg	Ala	Met	Arg	Cys	Ser	Leu	Phe	Arg	Thr	Lys	His	His	Ala	Arg	Leu	
				165						170					175		
30	Ser	Cys	Gly	Leu	Ile	Trp	Met	Ala	Ser	Val	Ser	Ala	Thr	Leu	Val	Pro	
				180					185					190			
	Phe	Thr	Ala	Val	His	Leu	Gln	His	Thr	Asp	Glu	Ala	Cys	Phe	Cys	Phe	
			195				200						205				
	Ala	Asp	Val	Arg	Glu	Val	Gln	Trp	Leu	Glu	Val	Thr	Leu	Gly	Phe	Ile	
			210				215						220				
35	Val	Pro	Phe	Ala	Ile	Ile	Gly	Leu	Cys	Tyr	Ser	Leu	Ile	Val	Arg	Val	
	225					230					235					240	
	Leu	Val	Arg	Ala	His	Arg	His	Arg	Gly	Leu	Arg	Pro	Arg	Arg	Gln	Lys	

45

[illegible]

(58) INFORMATION FOR SEQ ID NO:57:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

AAGGAATTCA CGGCCGGGGTG ATGCCATTCC C

31

(59) INFORMATION FOR SEQ ID NO:58:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

35 GGTGGATCCA TAAACACGGG CGTTGAGGAC

30

(60) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 960 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

ATGCCATTCC	CAAACTGCTC	AGCCCCCAGC	ACTGTGGTGG	CCACAGCTGT	GGGTGTCTTG	60
CTGGGGCTGG	AGTGTGGGCT	GGGTCTGCTG	GGCAACGCGG	TGGCGCTGTG	GACCTTCCTG	120
TTCCGGGTCA	GGGTGTGGAA	GCCGTACGCT	GTCTACCTGC	TCAACCTGGC	CCTGGCTGAC	180
CTGCTGTTGG	CTGCGTGCCT	GCCTTTCCTG	GCCGCCTTCT	ACCTGAGCCT	CCAGGCTTGG	240
CATCTGGGCC	GTGTGGGCTG	CTGGGCCCTG	CGCTTCCTGC	TGGACCTCAG	CCGCAGCGTG	300
GGGATGGCCT	TCCTGGCCGC	CGTGGCTTTG	GACCGGTACC	TCCGTGTGGT	CCACCCTCGG	360
CTTAAGGTCA	ACCTGCTGTC	TCCTCAGGCG	GCCCTGGGGG	TCTCGGGCCT	CGTCTGGCTC	420
CTGATGGTCG	CCCTCACCTG	CCCGGGCTTG	CTCATCTCTG	AGGCCGCCCA	GAACTCCACC	480
AGGTGCCACA	GTTTCTACTC	CAGGGCAGAC	GGCTCCTTCA	GCATCATCTG	GCAGGAAGCA	540
CTCTCCTGCC	TTCAGTTTGT	CCTCCCCTTT	GGCCTCATCG	TGTTCTGCAA	TGCAGGCATC	600
ATCAGGGCTC	TCCAGAAAAG	ACTCCGGGAG	CCTGAGAAAC	AGCCCCAAGCT	TCAGCGGGCC	660
CAGGCACTGG	TCACCTTGGT	GGTGGTGCTG	TTTGCTCTGT	GCTTCTGCCC	CTGCTTCCTG	720
GCCAGAGTCC	TGATGCACAT	CTTCCAGAAT	CTGGGGAGCT	GCAGGGCCCT	TTGTGCAGTG	780
GCTCATACCT	CGGATGTCAC	GGGCAGCCTC	ACCTACCTGC	ACAGTGTCGT	CAACCCCGTG	840
GTATACTGCT	TCTCCAGCCC	CACCTTCAGG	AGCTCCTATC	GGAGGGTCTT	CCACACCCTC	900
CGAGGCAAAG	GGCAGGCAGC	AGAGCCCCCA	GATTTCAACC	CCAGAGACTC	CTATTCCTGA	960

(61) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 319 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Met Pro Phe Pro Asn Cys Ser Ala Pro Ser Thr Val Val Ala Thr Ala

47

	1		5		10		15
	Val	Gly	Val	Leu	Gly	Leu	Glu
			20			25	
	Cys	Gly	Leu	Gly	Leu	Gly	Leu
			30				
	Ala	Val	Ala	Leu	Trp	Thr	Phe
			35			40	
	Leu	Phe	Arg	Val	Arg	Val	Trp
			45				
5	Tyr	Ala	Val	Tyr	Leu	Leu	Asn
			50			55	
	Leu	Ala	Leu	Ala	Asp	Leu	Leu
			60				
	Ala	Cys	Leu	Pro	Phe	Leu	Ala
			65			70	
	Phe	Tyr	Leu	Ser	Leu	Gln	Ala
			75			80	
10	His	Leu	Gly	Arg	Val	Gly	Cys
			85			90	
	Ala	Leu	Arg	Phe	Leu	Leu	Asp
			95				
	Ser	Arg	Ser	Val	Gly	Met	Ala
			100			105	
	Phe	Leu	Ala	Ala	Val	Ala	Leu
			110				
15	Tyr	Leu	Arg	Val	Val	His	Pro
			115			120	
	Arg	Leu	Lys	Val	Asn	Leu	Leu
			125				
	Gln	Ala	Ala	Leu	Gly	Val	Ser
			130			135	
	Gly	Leu	Val	Trp	Leu	Leu	Met
			140				
	Leu	Thr	Cys	Pro	Gly	Leu	Leu
			145			150	
	Ile	Ser	Glu	Ala	Ala	Gln	Asn
			155				
20	Arg	Cys	His	Ser	Phe	Tyr	Ser
			165			170	
	Ala	Asp	Gly	Ser	Phe	Ser	Ile
			175				
	Trp	Gln	Glu	Ala	Leu	Ser	Cys
			180			185	
	Phe	Val	Leu	Pro	Phe	Gly	Leu
			190				
25	Ile	Val	Phe	Cys	Asn	Ala	Gly
			195			200	
	Ile	Arg	Ala	Leu	Gln	Lys	Arg
			205				
	Arg	Glu	Pro	Glu	Lys	Gln	Pro
			210			215	
	Leu	Gln	Arg	Ala	Gln	Ala	Leu
			220				
	Thr	Leu	Val	Val	Val	Leu	Phe
			225			230	
	Ala	Leu	Cys	Phe	Leu	Pro	Cys
			235			240	
30	Ala	Arg	Val	Leu	Met	His	Ile
			245			250	
	Phe	Gln	Asn	Leu	Gly	Ser	Cys
			255				
	Leu	Cys	Ala	Val	Ala	His	Thr
			260			265	
	Val	Thr	Gly	Ser	Leu	Thr	Tyr
			270				
35	Leu	His	Ser	Val	Val	Asn	Pro
			275			280	
	Val	Tyr	Cys	Phe	Ser	Ser	Pro
			285				
	Phe	Arg	Ser	Ser	Tyr	Arg	Arg
			290			295	
	Val	Phe	His	Thr	Leu	Arg	Gly
			300				

Gln Ala Ala Glu Pro Pro Asp Phe Asn Pro Arg Asp Ser Tyr Ser
305 310 315

(62) INFORMATION FOR SEQ ID NO:61:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1143 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:
- | | | | | | | |
|------------|------------|------------|------------|------------|------------|------|
| ATGGAGGAAG | GTGGTGATTT | TGACAACTAC | TATGGGGCAG | ACAACCAGTC | TGAGTGTGAG | 60 |
| TACACAGACT | GGAAATCCTC | GGGGGCCCTC | ATCCCTGCCA | TCTACATGTT | GGTCTTCCTC | 120 |
| CTGGGCACCA | CGGGAAACGG | TCTGGTGCTC | TGGACCGTGT | TTCGGAGCAG | CCGGGAGAAG | 180 |
| AGGCGCTCAG | CTGATATCTT | CATTGCTAGC | CTGGCGGTGG | CTGACCTGAC | CTTCGTGGTG | 240 |
| 15 | ACGCTGCCCC | TGTGGGCTAC | CTACACGTAC | CGGGACTATG | ACTGGCCCTT | 300 |
| | TTCTGCAAGC | TCAGCAGCTA | CCTCATCTTC | GTCAACATGT | ACGCCAGCGT | 360 |
| | ACCGGCCTCA | GCTTCGACCG | CTACCTGGCC | ATCGTGAGGC | CAGTGGCCAA | 420 |
| | AGGCTGCGGG | TCAGCGGGGC | CGTGGCCACG | GCAGTTCCTT | GGGTGCTGGC | 480 |
| | GCCATGCCTG | TCATGGTGTT | ACGCACCACC | GGGGACTTGG | AGAACACCAC | 540 |
| 20 | TGCTACATGG | ACTACTCCAT | GGTGGCCACT | GTGAGCTCAG | AGTGGGCCTG | 600 |
| | CTTGGGGTCT | CGTCCACCAC | CGTGGGCTTT | GTGGTGCCCT | TCACCATCAT | 660 |
| | TACTTCTTCA | TCGCCCCAAC | CATCGCTGGC | CACTTCCGCA | AGGAACGCAT | 720 |
| | CGGAAGCGGC | GCCGGCTGCT | CAGCATCATC | GTGGTGCTGG | TGGTGACCTT | 780 |
| | TGGATGCCCT | ACCACCTGGT | GAAGACGCTG | TACATGCTGG | GCAGCCTGCT | 840 |
| 25 | TGTGACTTTG | ACCTCTTCCT | CATGAACATC | TTCCCTACT | GCACCTGCAT | 900 |
| | AACAGCTGCC | TCAACCCCTT | CCTCTATGCC | TTTTTCGACC | CCCGCTTCCG | 960 |
| | ACCTCCATGC | TCTGCTGTGG | CCAGAGCAGG | TGCGCAGGCA | CCTCCCACAG | 1020 |
| | GAGAAGTCAG | CCAGCTACTC | TTCGGGGCAC | AGCCAGGGGC | CCGGCCCCAA | 1080 |
| | GGTGAGAAAC | AGATGCACGA | GAAATCCATC | CCCTACAGCC | AGGAGACCCT | 1140 |
| 30 | TAG | | | | | 1143 |

(63) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

10 Met Glu Glu Gly Gly Asp Phe Asp Asn Tyr Tyr Gly Ala Asp Asn Gln
 1 5 10 15
 Ser Glu Cys Glu Tyr Thr Asp Trp Lys Ser Ser Gly Ala Leu Ile Pro
 20 25 30
 Ala Ile Tyr Met Leu Val Phe Leu Leu Gly Thr Thr Gly Asn Gly Leu
 35 40 45
 15 Val Leu Trp Thr Val Phe Arg Ser Ser Arg Glu Lys Arg Arg Ser Ala
 50 55 60
 Asp Ile Phe Ile Ala Ser Leu Ala Val Ala Asp Leu Thr Phe Val Val
 65 70 75 80
 20 Thr Leu Pro Leu Trp Ala Thr Tyr Thr Tyr Arg Asp Tyr Asp Trp Pro
 85 90 95
 Phe Gly Thr Phe Phe Cys Lys Leu Ser Ser Tyr Leu Ile Phe Val Asn
 100 105 110
 Met Tyr Ala Ser Val Phe Cys Leu Thr Gly Leu Ser Phe Asp Arg Tyr
 115 120 125
 25 Leu Ala Ile Val Arg Pro Val Ala Asn Ala Arg Leu Arg Leu Arg Val
 130 135 140
 Ser Gly Ala Val Ala Thr Ala Val Leu Trp Val Leu Ala Ala Leu Leu
 145 150 155 160
 30 Ala Met Pro Val Met Val Leu Arg Thr Thr Gly Asp Leu Glu Asn Thr
 165 170 175
 Thr Lys Val Gln Cys Tyr Met Asp Tyr Ser Met Val Ala Thr Val Ser
 180 185 190
 Ser Glu Trp Ala Trp Glu Val Gly Leu Gly Val Ser Ser Thr Thr Val
 195 200 205
 35 Gly Phe Val Val Pro Phe Thr Ile Met Leu Thr Cys Tyr Phe Phe Ile
 210 215 220
 Ala Gln Thr Ile Ala Gly His Phe Arg Lys Glu Arg Ile Glu Gly Leu

50

	225		230		235		240
	Arg Lys Arg Arg Arg	Leu Leu Ser Ile Ile	Val Val Leu Val Val	Thr			
		245		250		255	
5	Phe Ala Leu Cys Trp Met Pro Tyr His	Leu Val Lys Thr	Leu Tyr Met				
		260		265		270	
	Leu Gly Ser Leu Leu His Trp Pro Cys Asp Phe Asp	Leu Phe Leu Met					
		275		280		285	
	Asn Ile Phe Pro Tyr Cys Thr Cys Ile Ser Tyr Val	Asn Ser Cys Leu					
		290		295		300	
10	Asn Pro Phe Leu Tyr Ala Phe Phe Asp Pro Arg Phe Arg Gln Ala Cys						
		305		310		315	
	Thr Ser Met Leu Cys Cys Gly Gln Ser Arg Cys Ala Gly Thr Ser His						
		325		330		335	
15	Ser Ser Ser Gly Glu Lys Ser Ala Ser Tyr Ser Ser Gly His Ser Gln						
		340		345		350	
	Gly Pro Gly Pro Asn Met Gly Lys Gly Gly Glu Gln Met His Glu Lys						
		355		360		365	
	Ser Ile Pro Tyr Ser Gln Glu Thr Leu Val Val Asp						
		370		375		380	

20 (64) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGAGAATTCT GGTGACTCAC AGCCGGCACA G

31

(65) INFORMATION FOR SEQ ID NO:64:

30

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

35

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

GCCGGATCCA AGGAAAAGCA GCAATAAAAG G

31

(66) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1119 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

10 ATGAACTACC CGCTAACGCT GGAAATGGAC CTCGAGAACC TGGAGGACCT GTTCTGGGAA 60
CTGGACAGAT TGGACAAC TAACGACACC TCCCTGGTGG AAAATCATCT CTGCCCTGCC 120
ACAGAGGGTC CCCTCATGGC CTCCTTCAAG GCCGTGTTCTG TGCCCGTGGC CTACAGCCTC 180
ATCTTCCTCC TGGGCGTGAT CGGCAACGTC CTGGTGCTGG TGATCCTGGA GCGGCACCGG 240
CAGACACGCA GTTCCACGGA GACCTTCCTG TTCCACCTGG CCGTGGCCGA CCTCCTGCTG 300
15 GTCTTCATCT TGCCCTTTGC CGTGGCCGAG GGCTCTGTGG GCTGGGTCCT GGGGACCTTC 360
CTCTGCAAAA CTGTGATTGC CCTGCACAAA GTCAACTTCT ACTGCAGCAG CCTGCTCCTG 420
GCCTGCATCG CCGTGGACCG CTACCTGGCC ATTGTCCACG CCGTCCATGC CTACCGCCAC 480
CGCCGCCTCC TCTCCATCCA CATCACCTGT GGGACCATCT GGCTGGTGGG CTTCCTCCTT 540
GCCTTGCCAG AGATTCTCTT CGCCAAAGTC AGCCAAGGCC ATCACAACAA CTCCCTGCCA 600
20 CGTTGCACCT TCTCCAAGA GAACCAAGCA GAAACGCATG CCTGGTTCAC CTCCCGATTC 660
CTCTACCATG TGGCGGGATT CCTGCTGCCC ATGCTGGTGA TGGGCTGGTG CTACGTGGGG 720
GTAGTGACACA GGTGCGCCA GGCCAGCGG CGCCCTCAGC GGCAGAAGGC AGTCAGGGTG 780
GCCATCCTGG TGACAAGCAT CTTCTTCCTC TGCTGGTCAC CCTACCACAT CGTCATCTTC 840
CTGGACACCC TGGCGAGGCT GAAGGCCGTG GACAATACCT GCAAGCTGAA TGGCTCTCTC 900
25 CCCGTGGCCA TCACCATGTG TGAGTTCCTG GGCCTGGCCC ACTGCTGCCT CAACCCCATG 960
CTCTACACTT TCGCCGGCGT GAAGTTCCGC AGTGACCTGT CGCGGCTCCT GACCAAGCTG 1020
GGCTGTACCG GCCCTGCCTC CCTGTGCCAG CTCTTCCTTA GCTGGCGCAG GAGCAGTCTC 1080
TCTGAGTCAG AGAATGCCAC CTCTCTCACC ACGTTCTAG 1119

(67) INFORMATION FOR SEQ ID NO:66:

30 (i) SEQUENCE CHARACTERISTICS:

52

- (A) LENGTH: 372 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Met Asn Tyr Pro Leu Thr Leu Glu Met Asp Leu Glu Asn Leu Glu Asp
 1 5 10 15

Leu Phe Trp Glu Leu Asp Arg Leu Asp Asn Tyr Asn Asp Thr Ser Leu
 20 25 30

Val Glu Asn His Leu Cys Pro Ala Thr Glu Gly Pro Leu Met Ala Ser
 35 40 45

Phe Lys Ala Val Phe Val Pro Val Ala Tyr Ser Leu Ile Phe Leu Leu
 50 55 60

Gly Val Ile Gly Asn Val Leu Val Leu Val Ile Leu Glu Arg His Arg
 65 70 75 80

Gln Thr Arg Ser Ser Thr Glu Thr Phe Leu Phe His Leu Ala Val Ala
 85 90 95

Asp Leu Leu Leu Val Phe Ile Leu Pro Phe Ala Val Ala Glu Gly Ser
 100 105 110

Val Gly Trp Val Leu Gly Thr Phe Leu Cys Lys Thr Val Ile Ala Leu
 115 120 125

His Lys Val Asn Phe Tyr Cys Ser Ser Leu Leu Leu Ala Cys Ile Ala
 130 135 140

Val Asp Arg Tyr Leu Ala Ile Val His Ala Val His Ala Tyr Arg His
 145 150 155 160

Arg Arg Leu Leu Ser Ile His Ile Thr Cys Gly Thr Ile Trp Leu Val
 165 170 175

Gly Phe Leu Leu Ala Leu Pro Glu Ile Leu Phe Ala Lys Val Ser Gln
 180 185 190

Gly His His Asn Asn Ser Leu Pro Arg Cys Thr Phe Ser Gln Glu Asn
 195 200 205

Gln Ala Glu Thr His Ala Trp Phe Thr Ser Arg Phe Leu Tyr His Val
 210 215 220

Ala Gly Phe Leu Leu Pro Met Leu Val Met Gly Trp Cys Tyr Val Gly
 225 230 235 240

Val Val His Arg Leu Arg Gln Ala Gln Arg Arg Pro Gln Arg Gln Lys

[illegible]

(68) INFORMATION FOR SEQ ID NO:67:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

CAAAGCTTGA AAGCTGCACG GTGCAGAGAC

30

(69) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

35 GCGGATCCCG AGTCACACCC TGGCTGGGCC

30

(70) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1128 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

ATGGATGTGA CTTCCCAAGC CCGGGGCGTG GGCCTGGAGA TGTACCCAGG CACCGCGCAG 60
CCTGCGGCCC CCAACACCAC CTCCTCCGAG CTCAACCTGT CCCACCCGCT CCTGGGCACC 120
10 GCCCTGGCCA ATGGGACAGG TGAGCTCTCG GAGCACCAGC AGTACGTGAT CGGCCTGTTC 180
CTCTCGTGCC TCTACACCAT CTTCTCTTTC CCCATCGGCT TTGTGGGCAA CATCTGATC 240
CTGGTGGTGA ACATCAGCTT CCGCGAGAAG ATGACCATCC CCGACCTGTA CTTTCATCAAC 300
CTGGCGGTGG CGGACCTCAT CCTGGTGGCC GACTCCCTCA TTGAGGTGTT CAACCTGCAC 360
GAGCGGTACT ACGACATCGC CGTCTGTGC ACCTTCATGT CGCTCTTCCT GCAGGTCAAC 420
15 ATGTACAGCA GCGTCTTCTT CCTCACCTGG ATGAGCTTCG ACCGCTACAT CGCCCTGGCC 480
AGGGCCATGC GCTGCAGCCT GTTCCGCACC AAGCACCACG CCCGGCTGAG CTGTGGCCTC 540
ATCTGGATGG CATCCGTGTC AGCCACGCTG GTGCCCTTCA CCGCCGTGCA CCTGCAGCAC 600
ACCGACGAGG CCTGCTTCTG TTTCGCGGAT GTCCGGGAGG TGCAGTGGCT CGAGGTCACG 660
CTGGGCTTCA TCGTGCCCTT CGCCATCATC GGCCTGTGCT ACTCCCTCAT TGTCCGGGTG 720
20 CTGGTCAGGG CGCACC GGCA CCGTGGGCTG CGGCCCCGGC GGCAGAAGGC GCTCCGCATG 780
ATCCTCGCGG TGGTGCTGGT CTTCTTCGTC TGCTGGCTGC CGGAGAACGT CTTTCATCAGC 840
GTGCACCTCC TGCAGCGGAC GCAGCCTGGG GCCGCTCCCT GCAAGCAGTC TTTCCGCCAT 900
GCCCCACCCC TCACGGGCCA CATTGTCAAC CTCACCGCCT TCTCCAACAG CTGCCTAAAC 960
CCCCTCATCT ACAGCTTTCT CGGGGAGACC TTCAGGGACA AGCTGAGGCT GTACATTGAG 1020
25 CAGAAAACAA ATTTGCCGGC CCTGAACCGC TTCTGTCACG CTGCCCTGAA GGCCGTCATT 1080
CCAGACAGCA CCGAGCAGTC GGATGTGAGG TTCAGCAGTG CCGTGTAG 1128

(71) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 375 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

	Met	Asp	Val	Thr	Ser	Gln	Ala	Arg	Gly	Val	Gly	Leu	Glu	Met	Tyr	Pro	
	1				5					10					15		
5	Gly	Thr	Ala	Gln	Pro	Ala	Ala	Pro	Asn	Thr	Thr	Ser	Pro	Glu	Leu	Asn	
				20					25					30			
	Leu	Ser	His	Pro	Leu	Leu	Gly	Thr	Ala	Leu	Ala	Asn	Gly	Thr	Gly	Glu	
			35					40					45				
	Leu	Ser	Glu	His	Gln	Gln	Tyr	Val	Ile	Gly	Leu	Phe	Leu	Ser	Cys	Leu	
10			50				55					60					
	Tyr	Thr	Ile	Phe	Leu	Phe	Pro	Ile	Gly	Phe	Val	Gly	Asn	Ile	Leu	Ile	
	65					70					75				80		
	Leu	Val	Val	Asn	Ile	Ser	Phe	Arg	Glu	Lys	Met	Thr	Ile	Pro	Asp	Leu	
				85						90					95		
15	Tyr	Phe	Ile	Asn	Leu	Ala	Val	Ala	Asp	Leu	Ile	Leu	Val	Ala	Asp	Ser	
				100					105					110			
	Leu	Ile	Glu	Val	Phe	Asn	Leu	His	Glu	Arg	Tyr	Tyr	Asp	Ile	Ala	Val	
			115					120					125				
20	Leu	Cys	Thr	Phe	Met	Ser	Leu	Phe	Leu	Gln	Val	Asn	Met	Tyr	Ser	Ser	
		130					135						140				
	Val	Phe	Phe	Leu	Thr	Trp	Met	Ser	Phe	Asp	Arg	Tyr	Ile	Ala	Leu	Ala	
	145					150					155				160		
	Arg	Ala	Met	Arg	Cys	Ser	Leu	Phe	Arg	Thr	Lys	His	His	Ala	Arg	Leu	
				165						170					175		
25	Ser	Cys	Gly	Leu	Ile	Trp	Met	Ala	Ser	Val	Ser	Ala	Thr	Leu	Val	Pro	
				180				185						190			
	Phe	Thr	Ala	Val	His	Leu	Gln	His	Thr	Asp	Glu	Ala	Cys	Phe	Cys	Phe	
			195					200					205				
30	Ala	Asp	Val	Arg	Glu	Val	Gln	Trp	Leu	Glu	Val	Thr	Leu	Gly	Phe	Ile	
		210					215					220					
	Val	Pro	Phe	Ala	Ile	Ile	Gly	Leu	Cys	Tyr	Ser	Leu	Ile	Val	Arg	Val	
	225					230					235				240		
	Leu	Val	Arg	Ala	His	Arg	His	Arg	Gly	Leu	Arg	Pro	Arg	Arg	Gln	Lys	
				245						250					255		
35	Ala	Leu	Arg	Met	Ile	Leu	Ala	Val	Val	Leu	Val	Phe	Phe	Val	Cys	Trp	
				260				265							270		

56

Leu Pro Glu Asn Val Phe Ile Ser Val His Leu Leu Gln Arg Thr Gln
 275 280 285
 Pro Gly Ala Ala Pro Cys Lys Gln Ser Phe Arg His Ala His Pro Leu
 290 295 300
 5 Thr Gly His Ile Val Asn Leu Thr Ala Phe Ser Asn Ser Cys Leu Asn
 305 310 315 320
 Pro Leu Ile Tyr Ser Phe Leu Gly Glu Thr Phe Arg Asp Lys Leu Arg
 325 330 335
 10 Leu Tyr Ile Glu Gln Lys Thr Asn Leu Pro Ala Leu Asn Arg Phe Cys
 340 345 350
 His Ala Ala Leu Lys Ala Val Ile Pro Asp Ser Thr Glu Gln Ser Asp
 355 360 365
 Val Arg Phe Ser Ser Ala Val
 370 375

15 (72) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 20 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

ACAGAATTCC TGTGTGGTTT TACCGCCAG

30

(73) INFORMATION FOR SEQ ID NO:72:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 30 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CTCGGATCCA GGCAGAAGAG TCGCCTATGG

30

(74) INFORMATION FOR SEQ ID NO:73:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1137 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

```
ATGGACCTGG GGAAACCAAT GAAAAGCGTG CTGGTGGTGG CTCTCCTTGT CATTTTCCAG      60
5  GTATGCCTGT GTCAAGATGA GGTACCGGAC GATTACATCG GAGACAACAC CACAGTGGAC      120
TACACTTTGT TCGAGTCTTT GTGCTCCAAG AAGGACGTGC GGAACTTTAA AGCCTGGTTC      180
CTCCCTATCA TGTACTCCAT CATTGTGTTT GTGGGCCTAC TGGGCAATGG GCTGGTCGTG      240
TTGACCTATA TCTATTTCAA GAGGCTCAAG ACCATGACCG ATACCTACCT GCTCAACCTG      300
GCGGTGGCAG ACATCCTCTT CCTCCTGACC CTTCCCTTCT GGGCCTACAG CGCGGCCAAG      360
10 TCCTGGGTCT TCGGTGTCCA CTTTGTGAAG CTCATCTTTG CCATCTACAA GATGAGCTTC      420
TTCAGTGGCA TGCTCCTACT TCTTTGCATC AGCATTGACC GCTACGTGGC CATCGTCCAG      480
GCTGTCTCAG CTCACCGCCA CCGTGCCCGC GTCCTTCTCA TCAGCAAGCT GTCCTGTGTG      540
GGCATCTGGA TACTAGCCAC AGTGCTCTCC ATCCCAGAGC TCCTGTACAG TGACCTCCAG      600
AGGAGCAGCA GTGAGCAAGC GATGCGATGC TCTCTCATCA CAGAGCATGT GGAGGCCTTT      660
15 ATCACCATCC AGGTGGCCCA GATGGTGATC GGCTTTCTGG TCCCCCTGCT GGCCATGAGC      720
TTCTGTTACC TTGTCATCAT CCGCACCTG CTCCAGGCAC GCAACTTTGA GCGCAACAAG      780
GCCATCAAGG TGATCATCGC TGTGGTCGTG GTCTTCATAG TCTTCCAGCT GCCCTACAAT      840
GGGGTGGTCC TGGCCCAGAC GGTGGCCAAC TTCAACATCA CCAGTAGCAC CTGTGAGCTC      900
AGTAAGCAAC TCAACATCGC CTACGACGTC ACCTACAGCC TGGCCTGCGT CCGCTGCTGC      960
20 GTCAACCCCTT TCTTGACGC CTTTCATCGC GTCAAGTTCC GCAACGATCT CTTCAAGCTC      1020
TTCAAGGACC TGGGCTGCCT CAGCCAGGAG CAGCTCCGGC AGTGGTCTTC CTGTGCGCAC      1080
ATCCGGCGCT CCTCCATGAG TGTGGAGGCC GAGACCACCA CCACCTTCTC CCCATAG      1137
```

(75) INFORMATION FOR SEQ ID NO:74:

```
25  (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 378 amino acids
      (B) TYPE: amino acid
      (C) STRANDEDNESS:
      (D) TOPOLOGY: not relevant
```

(ii) MOLECULE TYPE: protein

```
30  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:
```

Met Asp Leu Gly Lys Pro Met Lys Ser Val Leu Val Val Ala Leu Leu
 1 5 10 15
 Val Ile Phe Gln Val Cys Leu Cys Cln Asp Glu Val Thr Asp Asp Tyr
 20 25 30
 5 Ile Gly Asp Asn Thr Thr Val Asp Tyr Thr Leu Phe Glu Ser Leu Cys
 35 40 45
 Ser Lys Lys Asp Val Arg Asn Phe Lys Ala Trp Phe Leu Pro Ile Met
 50 55 60
 10 Tyr Ser Ile Ile Cys Phe Val Gly Leu Leu Gly Asn Gly Leu Val Val
 65 70 75 80
 Leu Thr Tyr Ile Tyr Phe Lys Arg Leu Lys Thr Met Thr Asp Thr Tyr
 85 90 95
 Leu Leu Asn Leu Ala Val Ala Asp Ile Leu Phe Leu Leu Thr Leu Pro
 100 105 110
 15 Phe Trp Ala Tyr Ser Ala Ala Lys Ser Trp Val Phe Gly Val His Phe
 115 120 125
 Cys Lys Leu Ile Phe Ala Ile Tyr Lys Met Ser Phe Phe Ser Gly Met
 130 135 140
 20 Leu Leu Leu Leu Cys Ile Ser Ile Asp Arg Tyr Val Ala Ile Val Gln
 145 150 155 160
 Ala Val Ser Ala His Arg His Arg Ala Arg Val Leu Leu Ile Ser Lys
 165 170 175
 Leu Ser Cys Val Gly Ile Trp Ile Leu Ala Thr Val Leu Ser Ile Pro
 180 185 190
 25 Glu Leu Leu Tyr Ser Asp Leu Gln Arg Ser Ser Ser Glu Gln Ala Met
 195 200 205
 Arg Cys Ser Leu Ile Thr Glu His Val Glu Ala Phe Ile Thr Ile Gln
 210 215 220
 30 Val Ala Gln Met Val Ile Gly Phe Leu Val Pro Leu Leu Ala Met Ser
 225 230 235 240
 Phe Cys Tyr Leu Val Ile Ile Arg Thr Leu Leu Gln Ala Arg Asn Phe
 245 250 255
 Glu Arg Asn Lys Ala Ile Lys Val Ile Ile Ala Val Val Val Val Phe
 260 265 270
 35 Ile Val Phe Gln Leu Pro Tyr Asn Gly Val Val Leu Ala Gln Thr Val
 275 280 285

59

Ala Asn Phe Asn Ile Thr Ser Ser Thr Cys Glu Leu Ser Lys Gln Leu
 290 295 300

Asn Ile Ala Tyr Asp Val Thr Tyr Ser Leu Ala Cys Val Arg Cys Cys
 305 310 315 320

5 Val Asn Pro Phe Leu Tyr Ala Phe Ile Gly Val Lys Phe Arg Asn Asp
 325 330 335

Leu Phe Lys Leu Phe Lys Asp Leu Gly Cys Leu Ser Gln Glu Gln Leu
 340 345 350

10 Arg Gln Trp Ser Ser Cys Arg His Ile Arg Arg Ser Ser Met Ser Val
 355 360 365

Glu Ala Glu Thr Thr Thr Thr Phe Ser Pro
 370 375

(76) INFORMATION FOR SEQ ID NO:75:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CTGGAATTCA CCTGGACCAC CACCAATGGA TA

32

(77) INFORMATION FOR SEQ ID NO:76:

25 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

30 CTCGGATCCT GCAAAGTTTG TCATACAGTT

30

(78) INFORMATION FOR SEQ ID NO:77:

35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1085 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

ATGGATATAC AAATGGCAAA CAATTTTACT CCGCCCTCTG CAACTCCTCA GGGAAATGAC 60
 TGTGACCTCT ATGCACATCA CAGCACGGCC AGGATAGTAA TGCCTCTGCA TTACAGCCTC 120
 GTCTTCATCA TTGGGCTCGT GGGAAACTTA CTAGCCTTGG TCGTCATTGT TCAAAACAGG 180
 5 AAAAAAATCA ACTCTACCAC CCTCTATTCA ACAAATTTGG TGATTCTGA TATACTTTTT 240
 ACCACGGCTT TGCCTACACG AATAGCCTAC TATGCAATGG GCTTTGACTG GAGAATCGGA 300
 GATGCCTTGT GTAGGATAAC TCGCTAGTG TTTTACATCA ACACATATGC AGGTGTGAAC 360
 TTTATGACCT GCCTGAGTAT TGACCGCTTC ATTGCTGTGG TGCACCCTCT ACGCTACAAC 420
 AAGATAAAAA GGATTGAACA TGCAAAAGGC GTGTGCATAT TTGTCTGGAT TCTAGTATTT 480
 10 GCTCAGACAC TCCCACTCCT CATCAACCCT ATGTCAAAGC AGGAGGCTGA AAGGATTACA 540
 TGCATGGAGT ATCCAAACTT TGAAGAACT AAATCTCTTC CCTGGATTCT GCTTGGGGCA 600
 TGTTTCATAG GATATGTACT TCCAATTATA ATCATTCTCA TCTGCTATTC TCAGATCTGC 660
 TGCAAACTCT TCAGAACTGC CAAACAAAAC CCACTCACTG AGAAATCTGG TGTAACAAAA 720
 AAGGCTCTCA ACACAATTAT TCTTATTATT GTTGTGTTTG TTCTCTGTTT CACACCTTAC 780
 15 CATGTTGCAA TTATTCAACA TATGATTAAG AAGCTTCGTT TCTCTAATTT CCTGGAATGT 840
 AGCCAAAGAC ATTCGTTCCA GATTTCTCTG CACTTTACAG TATGCCTGAT GAACTTCAAT 900
 TGCTGCATGG ACCCTTTTAT CTACTTCTTT GCATGTAAAG GGTATAAGAG AAAGGTTATG 960
 AGGATGCTGA AACGGCAAGT CAGTGTATCG ATTTCTAGTG CTGTGAAGTC AGCCCCTGAA 1020
 GAAAATTCAC GTGAAATGAC AGAAACGCAG ATGATGATAC ATTCCAAGTC TTCAAATGGA 1080
 20 AAGTGA 1086

(79) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 361 amino acids
 (B) TYPE: amino acid
 25 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

30 Met Asp Ile Gln Met Ala Asn Asn Phe Thr Pro Pro Ser Ala Thr Pro
 1 5 10 15

61

Gln Gly Asn Asp Cys Asp Leu Tyr Ala His His Ser Thr Ala Arg Ile
 20 25 30
 Val Met Pro Leu His Tyr Ser Leu Val Phe Ile Ile Gly Leu Val Gly
 35 40 45
 5 Asn Leu Leu Ala Leu Val Val Ile Val Gln Asn Arg Lys Lys Ile Asn
 50 55 60
 Ser Thr Thr Leu Tyr Ser Thr Asn Leu Val Ile Ser Asp Ile Leu Phe
 65 70 75 80
 10 Thr Thr Ala Leu Pro Thr Arg Ile Ala Tyr Tyr Ala Met Gly Phe Asp
 85 90 95
 Trp Arg Ile Gly Asp Ala Leu Cys Arg Ile Thr Ala Leu Val Phe Tyr
 100 105 110
 Ile Asn Thr Tyr Ala Gly Val Asn Phe Met Thr Cys Leu Ser Ile Asp
 115 120 125
 15 Arg Phe Ile Ala Val Val His Pro Leu Arg Tyr Asn Lys Ile Lys Arg
 130 135 140
 Ile Glu His Ala Lys Gly Val Cys Ile Phe Val Trp Ile Leu Val Phe
 145 150 155 160
 20 Ala Gln Thr Leu Pro Leu Leu Ile Asn Pro Met Ser Lys Gln Glu Ala
 165 170 175
 Glu Arg Ile Thr Cys Met Glu Tyr Pro Asn Phe Glu Glu Thr Lys Ser
 180 185 190
 Leu Pro Trp Ile Leu Leu Gly Ala Cys Phe Ile Gly Tyr Val Leu Pro
 195 200 205
 25 Leu Ile Ile Ile Leu Ile Cys Tyr Ser Gln Ile Cys Cys Lys Leu Phe
 210 215 220
 Arg Thr Ala Lys Gln Asn Pro Leu Thr Glu Lys Ser Gly Val Asn Lys
 225 230 235 240
 30 Lys Ala Leu Asn Thr Ile Ile Leu Ile Ile Val Val Phe Val Leu Cys
 245 250 255
 Phe Thr Pro Tyr His Val Ala Ile Ile Gln His Met Ile Lys Lys Leu
 260 265 270
 Arg Phe Ser Asn Phe Leu Glu Cys Ser Gln Arg His Ser Phe Gln Ile
 275 280 285
 35 Ser Leu His Phe Thr Val Cys Leu Met Asn Phe Asn Cys Cys Met Asp
 290 295 300
 Pro Phe Ile Tyr Phe Phe Ala Cys Lys Gly Tyr Lys Arg Lys Val Met

305	310	315	320
Arg Met Leu Lys	Arg Gln Val Ser Val	Ser Ile Ser Ser	Ala Val Lys
	325	330	335
5 Ser Ala Pro Glu	Glu Asn Ser Arg	Glu Met Thr Glu	Thr Gln Met Met
	340	345	350
Ile His Ser Lys	Ser Ser Asn Gly	Lys	
	355	360	

(80) INFORMATION FOR SEQ ID NO:79:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

CTGGAATTCT CCTGCTCATC CAGCCATGCG G 31

(81) INFORMATION FOR SEQ ID NO:80:

- 20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

25 CCTGGATCCC CACCCCTACT GGGGCCTCAG 30

(82) INFORMATION FOR SEQ ID NO:81:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1446 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

ATGCGGTGGC TGTGGCCCT GGCTGTCTCT CTGCTGTGA TTTTGGCTGT GGGGCTAAGC 60
 35 AGGGTCTCTG GGGGTGCCCC CCTGCACCTG GGCAGGCACA GAGCCGAGAC CCAGGAGCAG 120

CAGAGCCGAT CCAAGAGGGG CACCGAGGAT GAGGAGGCCA AGGGCGTGCA GCAGTATGTG 180
CCTGAGGAGT GGGCGGAGTA CCCCCGGCCC ATTCACCTG CTGGCCTGCA GCCAACCAAG 240
CCCTTGGTGG CCACCAGCCC TAACCCCGAC AAGGATGGGG GCACCCAGAG CAGTGGGCAG 300
GAACTGAGGG GCAATCTGAC AGGGGCACCA GGGCAGAGGC TACAGATCCA GAACCCCTG 360
5 TATCCGGTGA CCGAGAGCTC CTACAGTGCC TATGCCATCA TGCTTCTGGC GCTGGTGGTG 420
TTTGGCGTGG GCATTGTGGG CAACCTGTGCG GTCATGTGCA TCGTGTGGCA CAGCTACTAC 480
CTGAAGAGCG CCTGGAACTC CATCCTTGCC AGCCTGGCCC TCTGGGATTT TCTGGTCTCTC 540
TTTTTCTGCC TCCCTATTGT CATCTTCAAC GAGATCACCA AGCAGAGGCT ACTGGGTGAC 600
GTTTCTTGTC GTGCCGTGCC CTTTCATGGAG GTCTCCTCTC TGGGAGTCAC GACTTTCAGC 660
10 CTCTGTGCCC TGGGCATTGA CCGCTTCCAC GTGGCCACCA GCACCTGCC CAAGGTGAGG 720
CCCATCGAGC GGTGCCAATC CATCCTGGCC AAGTTGGCTG TCATCTGGGT GGGCTCCATG 780
ACGCTGGCTG TGCCTGAGCT CCTGCTGTGG CAGCTGGCAC AGGAGCCTGC CCCACCATG 840
GGCACCTGG ACTCATGCAT CATGAAACCC TCAGCCAGCC TGCCCGAGTC CCTGTATTCA 900
CTGGTGATGA CCTACCAGAA CGCCCGCATG TGGTGGTACT TTGGCTGCTA CTTCTGCCTG 960
15 CCCATCCTCT TCACAGTCAC CTGCCAGCTG GTGACATGGC GGGTGCGAGG CCCTCCAGGG 1020
AGGAAGTCAG AGTGCAGGGC CAGCAAGCAC GAGCAGTGTG AGAGCCAGCT CAACAGCACC 1080
GTGGTGGGCC TGACCGTGGT CTACGCCTTC TGCACCCTCC CAGAGAACGT CTGCAACATC 1140
GTGGTGGCCT ACCTCTCCAC CGAGCTGACC CGCCAGACCC TGGACCTCCT GGGCCTCATC 1200
AACCAGTTCT CCACCTTCTT CAAGGGCGCC ATCACCCAG TGCTGCTCCT TTGCATCTGC 1260
20 AGGCCGCTGG GCCAGGCCTT CCTGGACTGC TGCTGCTGCT GCTGCTGTGA GGAGTGCGGC 1320
GGGGCTTCGG AGGCCTCTGC TGCCAATGGG TCGGACAACA AGCTCAAGAC CGAGGTGTCC 1380
TCTTCCATCT ACTTCCACAA GCCCAGGGAG TCACCCACAC TCCTGCCCCT GGGCACACCT 1440
TGCTGA 1446

(83) INFORMATION FOR SEQ ID NO:82:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 481 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant
- 30 (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

	Met	Arg	Trp	Leu	Trp	Pro	Leu	Ala	Val	Ser	Leu	Ala	Val	Ile	Leu	Ala	
	1				5					10					15		
5	Val	Gly	Leu	Ser	Arg	Val	Ser	Gly	Gly	Ala	Pro	Leu	His	Leu	Gly	Arg	
				20					25					30			
	His	Arg	Ala	Glu	Thr	Gln	Glu	Gln	Gln	Ser	Arg	Ser	Lys	Arg	Gly	Thr	
			35					40					45				
	Glu	Asp	Glu	Glu	Ala	Lys	Gly	Val	Gln	Gln	Tyr	Val	Pro	Glu	Glu	Trp	
		50					55					60					
10	Ala	Glu	Tyr	Pro	Arg	Pro	Ile	His	Pro	Ala	Gly	Leu	Gln	Pro	Thr	Lys	
	65					70					75					80	
	Pro	Leu	Val	Ala	Thr	Ser	Pro	Asn	Pro	Asp	Lys	Asp	Gly	Gly	Thr	Pro	
					85					90					95		
15	Asp	Ser	Gly	Gln	Glu	Leu	Arg	Gly	Asn	Leu	Thr	Gly	Ala	Pro	Gly	Gln	
				100					105					110			
	Arg	Leu	Gln	Ile	Gln	Asn	Pro	Leu	Tyr	Pro	Val	Thr	Glu	Ser	Ser	Tyr	
			115					120					125				
	Ser	Ala	Tyr	Ala	Ile	Met	Leu	Leu	Ala	Leu	Val	Val	Phe	Ala	Val	Gly	
		130					135					140					
20	Ile	Val	Gly	Asn	Leu	Ser	Val	Met	Cys	Ile	Val	Trp	His	Ser	Tyr	Tyr	
	145					150					155					160	
	Leu	Lys	Ser	Ala	Trp	Asn	Ser	Ile	Leu	Ala	Ser	Leu	Ala	Leu	Trp	Asp	
				165						170					175		
25	Phe	Leu	Val	Leu	Phe	Phe	Cys	Leu	Pro	Ile	Val	Ile	Phe	Asn	Glu	Ile	
				180					185					190			
	Thr	Lys	Gln	Arg	Leu	Leu	Gly	Asp	Val	Ser	Cys	Arg	Ala	Val	Pro	Phe	
			195					200						205			
	Met	Glu	Val	Ser	Ser	Leu	Gly	Val	Thr	Thr	Phe	Ser	Leu	Cys	Ala	Leu	
		210					215					220					
30	Gly	Ile	Asp	Arg	Phe	His	Val	Ala	Thr	Ser	Thr	Leu	Pro	Lys	Val	Arg	
	225					230					235					240	
	Pro	Ile	Glu	Arg	Cys	Gln	Ser	Ile	Leu	Ala	Lys	Leu	Ala	Val	Ile	Trp	
					245					250					255		
35	Val	Gly	Ser	Met	Thr	Leu	Ala	Val	Pro	Glu	Leu	Leu	Leu	Trp	Gln	Leu	
				260					265					270			
	Ala	Gln	Glu	Pro	Ala	Pro	Thr	Met	Gly	Thr	Leu	Asp	Ser	Cys	Ile	Met	

65

	275	280	285
	Lys Pro Ser Ala Ser Leu	Pro Glu Ser Leu Tyr	Ser Leu Val Met Thr
	290	295	300
5	Tyr Gln Asn Ala Arg Met Trp Trp Tyr Phe Gly Cys Tyr Phe Cys Leu		
	305	310	315 320
	Pro Ile Leu Phe Thr Val Thr Cys Gln Leu Val Thr Trp Arg Val Arg		
	325	330	335
	Gly Pro Pro Gly Arg Lys Ser Glu Cys Arg Ala Ser Lys His Glu Gln		
	340	345	350
10	Cys Glu Ser Gln Leu Asn Ser Thr Val Val Gly Leu Thr Val Val Tyr		
	355	360	365
	Ala Phe Cys Thr Leu Pro Glu Asn Val Cys Asn Ile Val Val Ala Tyr		
	370	375	380
15	Leu Ser Thr Glu Leu Thr Arg Gln Thr Leu Asp Leu Leu Gly Leu Ile		
	385	390	395 400
	Asn Gln Phe Ser Thr Phe Phe Lys Gly Ala Ile Thr Pro Val Leu Leu		
	405	410	415
	Leu Cys Ile Cys Arg Pro Leu Gly Gln Ala Phe Leu Asp Cys Cys Cys		
	420	425	430
20	Cys Cys Cys Cys Glu Glu Cys Gly Gly Ala Ser Glu Ala Ser Ala Ala		
	435	440	445
	Asn Gly Ser Asp Asn Lys Leu Lys Thr Glu Val Ser Ser Ser Ile Tyr		
	450	455	460
25	Phe His Lys Pro Arg Glu Ser Pro Pro Leu Leu Pro Leu Gly Thr Pro		
	465	470	475 480
	Cys		

(84) INFORMATION FOR SEQ ID NO:83:

- 30 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

ATGTGGAACG CGACGCCAG CG

(85) INFORMATION FOR SEQ ID NO:84:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TCATGTATTA ATACTAGATT CT

22

10 (86) INFORMATION FOR SEQ ID NO:85:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TACCATGTGG AACGCGACGC CCAGCGAAGA GCCGGGGT

38

(87) INFORMATION FOR SEQ ID NO:86:

- 20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CGGAATTCAT GTATTAATAC TAGATTCTGT CCAGGCCCG

39

(88) INFORMATION FOR SEQ ID NO:87:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1101 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

67

ATGTGGAACG CGACGCCAG CGAAGAGCCG GGGTTCAACC TCACACTGGC CGACCTGGAC 60
 TGGGATGCTT CCCCCGGCAA CGACTCGCTG GCGACGAGC TGCTGCAGCT CTTCCCCGCG 120
 CCGCTGCTGG CGGGCGTCAC AGCCACCTGC GTGGCACTCT TCGTGGTGGG TATCGCTGGC 180
 AACCTGCTCA CCATGCTGGT GGTGTCGCGC TTCCGCGAGC TGCGCACCAC CACCAACCTC 240
 5 TACCTGTCCA GCATGGCCTT CTCCGATCTG CTCATCTTCC TCTGCATGCC CCTGGACCTC 300
 GTTCGCCTCT GGCAGTACCG GCCCTGGAAC TTCGGCGACC TCCTCTGCAA ACTCTTCCAA 360
 TTCGTCACTG AGAGCTGCAC CTACGCCACG GTGCTCACCA TCACAGCGCT GAGCGTCGAG 420
 CGCTACTTCG CCATCTGCTT CCCACTCCGG GCCAAGGTGG TGGTCACCAA GGGGCGGGTG 480
 AAGCTGGTCA TCTTCGTCAT CTGGGCCGTG GCCTTCTGCA GCGCCGGGCC CATCTTCGTG 540
 10 CTAGTCGGGG TGGAGCACGA GAACGGCACC GACCCTTGGG ACACCAACGA GTGCCGCCCC 600
 ACCGAGTTTG CGGTGCGCTC TGGACTGCTC ACGGTCATGG TGTGGGTGTC CAGCATCTTC 660
 TTCTTCCTTC CTGTCTTCTG TCTCAGGTC CTCTACAGTC TCATCGGCAG GAAGCTGTGG 720
 CGGAGGAGGC GCGGCGATGC TGTCTGGGT GCCTCGCTCA GGGACCAGAA CCACAAGCAA 780
 ACCGTGAAAA TGCTGGCTGT AGTGGTGTTC GCCTTCATCC TCTGCTGGCT CCCCTTCCAC 840
 15 GTAGGGCGAT ATTTATTTTC CAAATCCTTT GAGCCTGGCT CTTGGAGAT TGCTCAGATC 900
 AGCCAGTACT GCAACCTCGT GTCCTTTGTC CTCTTCTACC TCAGTGCTGC CATCAACCCC 960
 ATTCTGTACA ACATCATGTC CAAGAAGTAC CGGTGGCAG GTTCAGACT TCTGGGATTC 1020
 GAACCCTTCT CCCAGAGAAA GCTCTCCACT CTGAAAGATG AAAGTTCTCG GGCCTGGACA 1080
 GAATCTAGTA TTAATACATG A 1101

20 (89) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 366 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

25 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Met Trp Asn Ala Thr Pro Ser Glu Glu Pro Gly Phe Asn Leu Thr Leu
 1 5 10 15

30 Ala Asp Leu Asp Trp Asp Ala Ser Pro Gly Asn Asp Ser Leu Gly Asp
 20 25 30

68

	Glu	Leu	Leu	Gln	Leu	Phe	Pro	Ala	Pro	Leu	Leu	Ala	Gly	Val	Thr	Ala		
		35						40					45					
	Thr	Cys	Val	Ala	Leu	Phe	Val	Val	Gly	Ile	Ala	Gly	Asn	Leu	Leu	Thr		
		50					55					60						
5	Met	Leu	Val	Val	Ser	Arg	Phe	Arg	Glu	Leu	Arg	Thr	Thr	Thr	Thr	Asn	Leu	
	65					70					75					80		
	Tyr	Leu	Ser	Ser	Met	Ala	Phe	Ser	Asp	Leu	Leu	Ile	Phe	Leu	Cys	Met		
					85					90					95			
10	Pro	Leu	Asp	Leu	Val	Arg	Leu	Trp	Gln	Tyr	Arg	Pro	Trp	Asn	Phe	Gly		
				100					105					110				
	Asp	Leu	Leu	Cys	Lys	Leu	Phe	Gln	Phe	Val	Ser	Glu	Ser	Cys	Thr	Tyr		
				115				120					125					
	Ala	Thr	Val	Leu	Thr	Ile	Thr	Ala	Leu	Ser	Val	Glu	Arg	Tyr	Phe	Ala		
				130			135					140						
15	Ile	Cys	Phe	Pro	Leu	Arg	Ala	Lys	Val	Val	Val	Thr	Lys	Gly	Arg	Val		
	145					150					155					160		
	Lys	Leu	Val	Ile	Phe	Val	Ile	Trp	Ala	Val	Ala	Phe	Cys	Ser	Ala	Gly		
					165				170						175			
20	Pro	Ile	Phe	Val	Leu	Val	Gly	Val	Glu	His	Glu	Asn	Gly	Thr	Asp	Pro		
				180					185					190				
	Trp	Asp	Thr	Asn	Glu	Cys	Arg	Pro	Thr	Glu	Phe	Ala	Val	Arg	Ser	Gly		
				195				200					205					
	Leu	Leu	Thr	Val	Met	Val	Trp	Val	Ser	Ser	Ile	Phe	Phe	Phe	Leu	Pro		
				210			215					220						
25	Val	Phe	Cys	Leu	Thr	Val	Leu	Tyr	Ser	Leu	Ile	Gly	Arg	Lys	Leu	Trp		
	225					230					235				240			
	Arg	Arg	Arg	Arg	Gly	Asp	Ala	Val	Val	Gly	Ala	Ser	Leu	Arg	Asp	Gln		
					245					250					255			
30	Asn	His	Lys	Gln	Thr	Val	Lys	Met	Leu	Ala	Val	Val	Val	Phe	Ala	Phe		
				260					265					270				
	Ile	Leu	Cys	Trp	Leu	Pro	Phe	His	Val	Gly	Arg	Tyr	Leu	Phe	Ser	Lys		
			275					280					285					
	Ser	Phe	Glu	Pro	Gly	Ser	Leu	Glu	Ile	Ala	Gln	Ile	Ser	Gln	Tyr	Cys		
		290					295					300						
35	Asn	Leu	Val	Ser	Phe	Val	Leu	Phe	Tyr	Leu	Ser	Ala	Ala	Ile	Asn	Pro		
	305					310					315				320			
	Ile	Leu	Tyr	Asn	Ile	Met	Ser	Lys	Lys	Tyr	Arg	Val	Ala	Val	Phe	Arg		

325 330 335

Leu Leu Gly Phe Glu Pro Phe Ser Gln Arg Lys Leu Ser Thr Leu Lys
340 345 350

5 Asp Glu Ser Ser Arg Ala Trp Thr Glu Ser Ser Ile Asn Thr
355 360 365

(90) INFORMATION FOR SEQ ID NO:89:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

GCAAGCTTGT GCCCTCACCA AGCCATGCGA GCC 33

15 (91) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CGGAATTCAG CAATGAGTTC CGACAGAAGC 30

(92) INFORMATION FOR SEQ ID NO:91:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1842 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

30 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

ATGCGAGCCC CGGGCGCGCT TCTCGCCCGC ATGTCGCGGC TACTGCTTCT GCTACTGCTC 60

AAGGTGTCTG CCTCTTCTGC CCTCGGGGTC GCCCTGCGT CCAGAAACGA AACTTGTCTG 120

GGGGAGAGCT GTGCACCTAC AGTGATCCAG CGCCGCGGCA GGGACGCCTG GGGACCGGGA 180

35 AATTCTGCAA GAGACGTTCT GCGAGCCCGA GCACCCAGGG AGGAGCAGGG GGCAGCGTTT 240

CTTGC GGGAC CCTCCTGGGA CCTGCCGGCG GCCCCGGGCC GTGACCCGGC TGCAGGCAGA 300
GGGGCGGAGG CGTCGGCAGC CGGACCCCCG GGACCTCCAA CCAGGCCACC TGGCCCCCTGG 360
AGGTGGAAAG GTGCTCGGGG TCAGGAGCCT TCTGAAACTT TGGGGAGAGG GAACCCACG 420
GCCCTCCAGC TCTTCCTTCA GATCTCAGAG GAGGAAGAGA AGGGTCCCAG AGGCGCTGGC 480
5 ATTTCCGGGC GTAGCCAGGA GCAGAGTGTG AAGACAGTCC CCGGAGCCAG CGATCTTTTT 540
TACTGGCCAA GGAGAGCCGG GAAACTCCAG GGTTCACACC ACAAGCCCCT GTCCAAGACG 600
GCCAATGGAC TGGCGGGGCA CGAAGGGTGG ACAATTGCAC TCCGGGGCCG GGCCTGGCC 660
CAGAAATGGAT CCTTGGGTGA AGGAATCCAT GAGCCTGGGG GTCCCCGCCG GGGAAACAGC 720
ACGAACCGGC GTGTGAGACT GAAGAACCCC TTCTACCCGC TGACCCAGGA GTCCTATGGA 780
10 GCCTACGCGG TCATGTGTCT GTCCGTGGTG ATCTTCGGGA CCGGCATCAT TGGCAACCTG 840
GCGGTGATGA GCATCGTGTG CCACAACCTAC TACATGCGGA GCATCTCCAA CTCCCTCTTG 900
GCCAACCTGG CCTTCTGGGA CTTTCTCATC ATCTTCTTCT GCCTTCCGCT GGTCTCTTC 960
CACGAGCTGA CCAAGAAGTG GCTGCTGGAG GACTTCTCCT GCAAGATCGT GCCCTATATA 1020
GAGGTCGCTT CTCTGGGAGT CACCACTTTC ACCTTATGTG CTCTGTGCAT AGACCGCTTC 1080
15 CGTGCTGCCA CCAACGTACA GATGTACTAC GAAATGATCG AAAACTGTTC CTCAACAACT 1140
GCCAACTTG CTGTTATATG GGTGGGAGCT CTATTGTTAG CACTTCCAGA AGTTGTTCTC 1200
CGCCAGCTGA GCAAGGAGGA TTTGGGGTTT AGTGGCCGAG CTCCGGCAGA AAGGTGCATT 1260
ATTAAGATCT CTCCTGATTT ACCAGACACC ATCTATGTTT TAGCCCTCAC CTACGACAGT 1320
GCGAGACTGT GGTGGTATTT TGGCTGTTAC TTTTGTTCGC CCACGCTTTT CACCATCACC 1380
20 TGCTCTCTAG TGA CTGCGAG GAAAATCCGC AAAGCAGAGA AAGCCTGTAC CCGAGGGAAT 1440
AAACGGCAGA TTCAACTAGA GAGTCAGATG AACTGTACAG TAGTGGCACT GACCATTTTA 1500
TATGGATTTT GCATTATTCC TGAAAATATC TGCAACATTG TTACTGCCTA CATGGCTACA 1560
GGGGTTTCAC AGCAGACAAT GGACCTCCTT AATATCATCA GCCAGTTCCT TTTGTTCTTT 1620
AAGTCCTGTG TCACCCAGT CCTCCTTTTC TGTCTCTGCA AACCCTTCAG TCGGGCCTTC 1680
25 ATGGAGTGCT GCTGCTGTTG CTGTGAGGAA TGCATTGAGA AGTCTTCAAC GGTGACCAGT 1740
GATGACAATG ACAACGAGTA CACCACGGAA CTCGAACCTC CGCCTTTCAG TACCATACGC 1800
CGTGAAATGT CCACTTTTGC TCTGTGCGGA ACTCATTGCT GA 1842

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 613 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

5

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

	Met	Arg	Ala	Pro	Gly	Ala	Leu	Leu	Ala	Arg	Met	Ser	Arg	Leu	Leu	Leu	
	1				5					10					15		
10	Leu	Leu	Leu	Leu	Lys	Val	Ser	Ala	Ser	Ser	Ala	Leu	Gly	Val	Ala	Pro	
				20					25					30			
	Ala	Ser	Arg	Asn	Glu	Thr	Cys	Leu	Gly	Glu	Ser	Cys	Ala	Pro	Thr	Val	
			35					40					45				
15	Ile	Gln	Arg	Arg	Gly	Arg	Asp	Ala	Trp	Gly	Pro	Gly	Asn	Ser	Ala	Arg	
	50						55					60					
	Asp	Val	Leu	Arg	Ala	Arg	Ala	Pro	Arg	Glu	Glu	Gln	Gly	Ala	Ala	Phe	
	65					70				75						80	
	Leu	Ala	Gly	Pro	Ser	Trp	Asp	Leu	Pro	Ala	Ala	Pro	Gly	Arg	Asp	Pro	
					85					90					95		
20	Ala	Ala	Gly	Arg	Gly	Ala	Glu	Ala	Ser	Ala	Ala	Gly	Pro	Pro	Gly	Pro	
				100					105					110			
	Pro	Thr	Arg	Pro	Pro	Gly	Pro	Trp	Arg	Trp	Lys	Gly	Ala	Arg	Gly	Gln	
			115					120					125				
25	Glu	Pro	Ser	Glu	Thr	Leu	Gly	Arg	Gly	Asn	Pro	Thr	Ala	Leu	Gln	Leu	
	130						135					140					
	Phe	Leu	Gln	Ile	Ser	Glu	Glu	Glu	Glu	Lys	Gly	Pro	Arg	Gly	Ala	Gly	
	145					150				155						160	
	Ile	Ser	Gly	Arg	Ser	Gln	Glu	Gln	Ser	Val	Lys	Thr	Val	Pro	Gly	Ala	
				165					170						175		
30	Ser	Asp	Leu	Phe	Tyr	Trp	Pro	Arg	Arg	Ala	Gly	Lys	Leu	Gln	Gly	Ser	
				180				185						190			
	His	His	Lys	Pro	Leu	Ser	Lys	Thr	Ala	Asn	Gly	Leu	Ala	Gly	His	Glu	
			195					200					205				
35	Gly	Trp	Thr	Ile	Ala	Leu	Pro	Gly	Arg	Ala	Leu	Ala	Gln	Asn	Gly	Ser	
	210						215					220					
	Leu	Gly	Glu	Gly	Ile	His	Glu	Pro	Gly	Gly	Pro	Arg	Gly	Asn	Ser		
	225				230					235					240		

	Thr	Asn	Arg	Arg	Val	Arg	Leu	Lys	Asn	Pro	Phe	Tyr	Pro	Leu	Thr	Gln	
					245					250					255		
	Glu	Ser	Tyr	Gly	Ala	Tyr	Ala	Val	Met	Cys	Leu	Ser	Val	Val	Ile	Phe	
				260					265					270			
5	Gly	Thr	Gly	Ile	Ile	Gly	Asn	Leu	Ala	Val	Met	Ser	Ile	Val	Cys	His	
			275					280					285				
	Asn	Tyr	Tyr	Met	Arg	Ser	Ile	Ser	Asn	Ser	Leu	Leu	Ala	Asn	Leu	Ala	
		290					295					300					
10	Phe	Trp	Asp	Phe	Leu	Ile	Ile	Phe	Phe	Cys	Leu	Pro	Leu	Val	Ile	Phe	
	305					310					315				320		
	His	Glu	Leu	Thr	Lys	Lys	Trp	Leu	Leu	Glu	Asp	Phe	Ser	Cys	Lys	Ile	
					325					330					335		
	Val	Pro	Tyr	Ile	Glu	Val	Ala	Ser	Leu	Gly	Val	Thr	Thr	Phe	Thr	Leu	
				340					345					350			
15	Cys	Ala	Leu	Cys	Ile	Asp	Arg	Phe	Arg	Ala	Ala	Thr	Asn	Val	Gln	Met	
		355						360					365				
	Tyr	Tyr	Glu	Met	Ile	Glu	Asn	Cys	Ser	Ser	Thr	Thr	Ala	Lys	Leu	Ala	
		370					375					380					
20	Val	Ile	Trp	Val	Gly	Ala	Leu	Leu	Leu	Ala	Leu	Pro	Glu	Val	Val	Leu	
	385				390						395				400		
	Arg	Gln	Leu	Ser	Lys	Glu	Asp	Leu	Gly	Phe	Ser	Gly	Arg	Ala	Pro	Ala	
				405					410					415			
	Glu	Arg	Cys	Ile	Ile	Lys	Ile	Ser	Pro	Asp	Leu	Pro	Asp	Thr	Ile	Tyr	
			420					425					430				
25	Val	Leu	Ala	Leu	Thr	Tyr	Asp	Ser	Ala	Arg	Leu	Trp	Trp	Tyr	Phe	Gly	
		435					440					445					
	Cys	Tyr	Phe	Cys	Leu	Pro	Thr	Leu	Phe	Thr	Ile	Thr	Cys	Ser	Leu	Val	
		450					455					460					
30	Thr	Ala	Arg	Lys	Ile	Arg	Lys	Ala	Glu	Lys	Ala	Cys	Thr	Arg	Gly	Asn	
	465				470					475					480		
	Lys	Arg	Gln	Ile	Gln	Leu	Glu	Ser	Gln	Met	Asn	Cys	Thr	Val	Val	Ala	
				485					490					495			
	Leu	Thr	Ile	Leu	Tyr	Gly	Phe	Cys	Ile	Ile	Pro	Glu	Asn	Ile	Cys	Asn	
		500						505					510				
35	Ile	Val	Thr	Ala	Tyr	Met	Ala	Thr	Gly	Val	Ser	Gln	Gln	Thr	Met	Asp	
		515					520					525					
	Leu	Leu	Asn	Ile	Ile	Ser	Gln	Phe	Leu	Leu	Phe	Phe	Lys	Ser	Cys	Val	

73

530 535 540

Thr Pro Val Leu Leu Phe Cys Leu Cys Lys Pro Phe Ser Arg Ala Phe
 545 550 555 560

Met Glu Cys Cys Cys Cys Cys Glu Glu Cys Ile Gln Lys Ser Ser
 5 565 570 575

Thr Val Thr Ser Asp Asp Asn Asp Asn Glu Tyr Thr Thr Glu Leu Glu
 580 585 590

Leu Ser Pro Phe Ser Thr Ile Arg Arg Glu Met Ser Thr Phe Ala Ser
 595 600 605

10 Val Gly Thr His Cys
 610

(94) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 base pairs
 15 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

20 CAGAATTCAG AGAAAAAAG TGAATATGGT TTTT

34

(95) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 25 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

TTGGATCCCT GGTGCATAAC AATTGAAAGA AT

32

30 (96) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1248 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ATGGTTTTTG CTCACAGAAT GGATAACAGC AAGCCACATT TGATTATTCC TACACTTCTG 60
 GTGCCCCCTCC AAAACCGCAG CTGCACTGAA ACAGCCACAC CTCTGCCAAG CCAATACCTG 120
 ATGGAATTAA GTGAGGAGCA CAGTTGGATG AGCAACCAAA CAGACCTTCA CTATGTGCTG 180
 5 AAACCCGGGG AAGTGGCCAC AGCCAGCATC TTCTTTGGGA TTCTGTGGTT GTTTTCTATC 240
 TTCGGCAATT CCCTGGTTTG TTTGGTCATC CATAGGAGTA GGAGGACTCA GTCTACCACC 300
 AACTACTTTG TGGTCTCCAT GGCATGTGCT GACCTTCTCA TCAGCGTTGC CAGCACGCCT 360
 TTCGTCTGTC TCCAGTTCAC CACTGGAAGG TGGACGCTGG GTAGTGCAAC GTGCAAGGTT 420
 GTGCGATATT TTCAATATCT CACTCCAGGT GTCCAGATCT ACGTTCTCCT CTCCATCTGC 480
 10 ATAGACCGGT TCTACACCAT CGTCTATCCT CTGAGCTTCA AGGTGTCCAG AGAAAAAGCC 540
 AAGAAAATGA TTGCGGCATC GTGGATCTTT GATGCAGGCT TTGTGACCCC TGTGCTCTTT 600
 TTCTATGGCT CCAACTGGGA CAGTCATTGT AACTATTTCC TCCCCTCCTC TTGGGAAGGC 660
 ACTGCCTACA CTGTCATCCA CTTCTTGGTG GGCTTTGTGA TTCCATCTGT CCTCATAATT 720
 TTATTTTACC AAAAGGTCAT AAAATATATT TGGAGAATAG GCACAGATGG CCGAACGGTG 780
 15 AGGAGGACAA TGAACATTGT CCCTCGGACA AAAGTGAAAA CTATCAAGAT GTTCCTCATT 840
 TTAAATCTGT TGTTTTTGCT CTCCTGGCTG CCTTTTCATG TAGCTCAGCT ATGGCACCCC 900
 CATGAACAAG ACTATAAGAA AAGTTCCTT GTTTTCACAG CTATCACATG GATATCCTTT 960
 AGTTCTTCAG CCTCTAAACC TACTCTGTAT TCAATTTATA ATGCCAATTT TCGGAGAGGG 1020
 ATGAAAGAGA CTTTTTG CAT GTCCTCTATG AAATGTTACC GAAGCAATGC CTATACTATC 1080
 20 ACAACAAGTT CAAGGATGGC CAAAAA AAC TACGTTGGCA TTTCAGAAAT CCCTTCCATG 1140
 GCCAAACTA TTACCAAAGA CTCGATCTAT GACTCATTG ACAGAGAAGC CAAGGAAAAA 1200
 AAGCTTGCTT GGCCCATTA CTCAAATCCA CCAAATACTT TTGTCTAA 1248

(97) INFORMATION FOR SEQ ID NO:96:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 415 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

75

	Met	Val	Phe	Ala	His	Arg	Met	Asp	Asn	Ser	Lys	Pro	His	Leu	Ile	Ile	
	1				5					10					15		
	Pro	Thr	Leu	Leu	Val	Pro	Leu	Gln	Asn	Arg	Ser	Cys	Thr	Glu	Thr	Ala	
				20					25					30			
5	Thr	Pro	Leu	Pro	Ser	Gln	Tyr	Leu	Met	Glu	Leu	Ser	Glu	Glu	His	Ser	
			35					40					45				
	Trp	Met	Ser	Asn	Gln	Thr	Asp	Leu	His	Tyr	Val	Leu	Lys	Pro	Gly	Glu	
		50					55					60					
10	Val	Ala	Thr	Ala	Ser	Ile	Phe	Phe	Gly	Ile	Leu	Trp	Leu	Phe	Ser	Ile	
	65					70					75					80	
	Phe	Gly	Asn	Ser	Leu	Val	Cys	Leu	Val	Ile	His	Arg	Ser	Arg	Arg	Thr	
					85					90					95		
	Gln	Ser	Thr	Thr	Asn	Tyr	Phe	Val	Val	Ser	Met	Ala	Cys	Ala	Asp	Leu	
				100					105						110		
15	Leu	Ile	Ser	Val	Ala	Ser	Thr	Pro	Phe	Val	Leu	Leu	Gln	Phe	Thr	Thr	
			115					120					125				
	Gly	Arg	Trp	Thr	Leu	Gly	Ser	Ala	Thr	Cys	Lys	Val	Val	Arg	Tyr	Phe	
		130					135					140					
20	Gln	Tyr	Leu	Thr	Pro	Gly	Val	Gln	Ile	Tyr	Val	Leu	Leu	Ser	Ile	Cys	
	145					150					155					160	
	Ile	Asp	Arg	Phe	Tyr	Thr	Ile	Val	Tyr	Pro	Leu	Ser	Phe	Lys	Val	Ser	
				165					170						175		
	Arg	Glu	Lys	Ala	Lys	Lys	Met	Ile	Ala	Ala	Ser	Trp	Ile	Phe	Asp	Ala	
				180					185						190		
25	Gly	Phe	Val	Thr	Pro	Val	Leu	Phe	Phe	Tyr	Gly	Ser	Asn	Trp	Asp	Ser	
			195					200					205				
	His	Cys	Asn	Tyr	Phe	Leu	Pro	Ser	Ser	Trp	Glu	Gly	Thr	Ala	Tyr	Thr	
		210					215					220					
30	Val	Ile	His	Phe	Leu	Val	Gly	Phe	Val	Ile	Pro	Ser	Val	Leu	Ile	Ile	
	225					230					235					240	
	Leu	Phe	Tyr	Gln	Lys	Val	Ile	Lys	Tyr	Ile	Trp	Arg	Ile	Gly	Thr	Asp	
				245						250					255		
	Gly	Arg	Thr	Val	Arg	Arg	Thr	Met	Asn	Ile	Val	Pro	Arg	Thr	Lys	Val	
				260					265						270		
35	Lys	Thr	Ile	Lys	Met	Phe	Leu	Ile	Leu	Asn	Leu	Leu	Phe	Leu	Leu	Ser	
			275					280					285				

76

Trp Leu Pro Phe His Val Ala Gln Leu Trp His Pro His Glu Gln Asp
 290 295 300
 Tyr Lys Lys Ser Ser Leu Val Phe Thr Ala Ile Thr Trp Ile Ser Phe
 305 310 315 320
 5 Ser Ser Ser Ala Ser Lys Pro Thr Leu Tyr Ser Ile Tyr Asn Ala Asn
 325 330 335
 Phe Arg Arg Gly Met Lys Glu Thr Phe Cys Met Ser Ser Met Lys Cys
 340 345 350
 10 Tyr Arg Ser Asn Ala Tyr Thr Ile Thr Thr Ser Ser Arg Met Ala Lys
 355 360 365
 Lys Asn Tyr Val Gly Ile Ser Glu Ile Pro Ser Met Ala Lys Thr Ile
 370 375 380
 Thr Lys Asp Ser Ile Tyr Asp Ser Phe Asp Arg Glu Ala Lys Glu Lys
 385 390 395 400
 15 Lys Leu Ala Trp Pro Ile Asn Ser Asn Pro Pro Asn Thr Phe Val
 405 410 415

(98) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

25 GGAAAGCTTA ACGATCCCCA GGAGCAACAT

30

(99) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

CTGGGATCCT ACGAGAGCAT TTTTCACACA G

31

35 (100) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1842 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

	ATGGGGCCCA	CCCTAGCGGT	TCCCACCCCC	TATGGCTGTA	TTGGCTGTAA	GCTACCCAG	60
	CCAGAATACC	CACCGGCTCT	AATCATCTTT	ATGTTCTGCG	CGATGGTTAT	CACCATCGTT	120
	GTAGACCTAA	TCGGCAACTC	CATGGTCATT	TTGGCTGTGA	CGAAGAACAA	GAAGCTCCGG	180
10	AATTCTGGCA	ACATCTTCGT	GGTCAGTCTC	TCTGTGGCCG	ATATGCTGGT	GGCCATCTAC	240
	CCATACCCTT	TGATGCTGCA	TGCCATGTCC	ATTGGGGGCT	GGGATCTGAG	CCAGTTACAG	300
	TGCCAGATGG	TCGGGTTCAT	CACAGGGCTG	AGTGTGGTCG	GCTCCATCTT	CAACATCGTG	360
	GCAATCGCTA	TCAACCGTTA	CTGCTACATC	TGCCACAGCC	TCCAGTACGA	ACGGATCTTC	420
	AGTGTGCGCA	ATACCTGCAT	CTACCTGGTC	ATCACCTGGA	TCATGACCGT	CCTGGCTGTC	480
15	CTGCCCAACA	TGTACATTGG	CACCATCGAG	TACGATCCTC	GCACCTACAC	CTGCATCTTC	540
	AACTATCTGA	ACAACCCTGT	CTTCACTGTT	ACCATCGTCT	GCATCCACTT	CGTCCTCCCT	600
	CTCCTCATCG	TGGGTTTCTG	CTACGTGAGG	ATCTGGACCA	AAGTGCTGGC	GGCCCGTGAC	660
	CCTGCAGGGC	AGAATCCTGA	CAACCAACTT	GCTGAGGTTC	GCAATTTTCT	AACCATGTTT	720
	GTGATCTTCC	TCCTCTTTGC	AGTGTGCTGG	TGCCCTATCA	ACGTGCTCAC	TGTCTTGGTG	780
20	GCTGTCAGTC	CGAAGGAGAT	GGCAGGCAAG	ATCCCCAACT	GGCTTTATCT	TGCAGCCTAC	840
	TTCATAGCCT	ACTTCAACAG	CTGCCTCAAC	GCTGTGATCT	ACGGGCTCCT	CAATGAGAAT	900
	TTCCGAAGAG	AATACTGGAC	CATCTTCCAT	GCTATGCGGC	ACCCTATCAT	ATTCTTCCCT	960
	GGCCTCATCA	GTGATATTCT	TGAGATGCAG	GAGGCCCCGT	CCCTGGCCCCG	CGCCCGTGCC	1020
	CATGCTCGCG	ACCAAGCTCG	TGAACAAGAC	CGTGCCCATG	CCTGTCCTGC	TGTGGAGGAA	1080
25	ACCCCGATGA	ATGTCCGGAA	TGTTCCATTA	CCTGGTGATG	CTGCAGCTGG	CCACCCCGAC	1140
	CGTGCTCTCT	GCCACCCTAA	GCCCCATTCC	AGATCCTCCT	CTGCCTATCG	CAAATCTGCC	1200
	TCTACCCACC	ACAAGTCTGT	CTTTAGCCAC	TCCAAGGCTG	CCTCTGGTCA	CCTCAAGCCT	1260
	GTCTCTGGCC	ACTCCAAGCC	TGCCTCTGGT	CACCCCAAGT	CTGCCACTGT	CTACCCTAAG	1320
	CCTGCCTCTG	TCCATTTCAA	GGGTGACTCT	GTCCATTTC	AGGGTGACTC	TGTCCATTTT	1380

AAGCCTGACT CTGTTTCATTT CAAGCCTGCT TCCAGCAACC CCAAGCCCAT CACTGGCCAC 1440
 CATGTCTCTG CTGGCAGCCA CTCCAAGTCT GCCTTCAGTG CTGCCACCAG CCACCCTAAA 1500
 CCCATCAAGC CAGCTACCAG CCATGCTGAG CCCACCACTG CTGACTATCC CAAGCCTGCC 1560
 ACTACCAGCC ACCCTAAGCC CGCTGCTGCT GACAACCCTG AGCTCTCTGC CTCCCATTGC 1620
 5 CCCGAGATCC CTGCCATTGC CCACCCTGTG TCTGACGACA GTGACCTCCC TGAGTCGGCC 1680
 TCTAGCCCTG CCGCTGGGCC CACCAAGCCT GCTGCCAGCC AGCTGGAGTC TGACACCATC 1740
 GCTGACCTTC CTGACCCTAC TGTAGTCACT ACCAGTACCA ATGATTACCA TGATGTCGTG 1800
 GTTGTGTGATG TTGAAGATGA TCCTGATGAA ATGGCTGTGT GA 1842

(101) INFORMATION FOR SEQ ID NO:100:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 613 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant
- 15 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:
- | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| Met | Gly | Pro | Thr | Leu | Ala | Val | Pro | Thr | Pro | Tyr | Gly | Cys | Ile | Gly | Cys | 1 | 5 | 10 | 15 |
| Lys | Leu | Pro | Gln | Pro | Glu | Tyr | Pro | Pro | Ala | Leu | Ile | Ile | Phe | Met | Phe | 20 | 25 | 30 | |
| Cys | Ala | Met | Val | Ile | Thr | Ile | Val | Val | Asp | Leu | Ile | Gly | Asn | Ser | Met | 35 | 40 | 45 | |
| Val | Ile | Leu | Ala | Val | Thr | Lys | Asn | Lys | Lys | Leu | Arg | Asn | Ser | Gly | Asn | 50 | 55 | 60 | |
| Ile | Phe | Val | Val | Ser | Leu | Ser | Val | Ala | Asp | Met | Leu | Val | Ala | Ile | Tyr | 65 | 70 | 75 | 80 |
| Pro | Tyr | Pro | Leu | Met | Leu | His | Ala | Met | Ser | Ile | Gly | Gly | Trp | Asp | Leu | 85 | 90 | 95 | |
| Ser | Gln | Leu | Gln | Cys | Gln | Met | Val | Gly | Phe | Ile | Thr | Gly | Leu | Ser | Val | 100 | 105 | 110 | |
| Val | Gly | Ser | Ile | Phe | Asn | Ile | Val | Ala | Ile | Ala | Ile | Asn | Arg | Tyr | Cys | 115 | 120 | 125 | |
| Tyr | Ile | Cys | His | Ser | Leu | Gln | Tyr | Glu | Arg | Ile | Phe | Ser | Val | Arg | Asn | 130 | 135 | 140 | |

Thr Cys Ile Tyr Leu Val Ile Thr Trp Ile Met Thr Val Leu Ala Val
 145 150 155 160
 Leu Pro Asn Met Tyr Ile Gly Thr Ile Glu Tyr Asp Pro Arg Thr Tyr
 165 170 175
 5 Thr Cys Ile Phe Asn Tyr Leu Asn Asn Pro Val Phe Thr Val Thr Ile
 180 185 190
 Val Cys Ile His Phe Val Leu Pro Leu Leu Ile Val Gly Phe Cys Tyr
 195 200 205
 10 Val Arg Ile Trp Thr Lys Val Leu Ala Ala Arg Asp Pro Ala Gly Gln
 210 215 220
 Asn Pro Asp Asn Gln Leu Ala Glu Val Arg Asn Phe Leu Thr Met Phe
 225 230 235 240
 Val Ile Phe Leu Leu Phe Ala Val Cys Trp Cys Pro Ile Asn Val Leu
 245 250 255
 15 Thr Val Leu Val Ala Val Ser Pro Lys Glu Met Ala Gly Lys Ile Pro
 260 265 270
 Asn Trp Leu Tyr Leu Ala Ala Tyr Phe Ile Ala Tyr Phe Asn Ser Cys
 275 280 285
 20 Leu Asn Ala Val Ile Tyr Gly Leu Leu Asn Glu Asn Phe Arg Arg Glu
 290 295 300
 Tyr Trp Thr Ile Phe His Ala Met Arg His Pro Ile Ile Phe Phe Pro
 305 310 315 320
 Gly Leu Ile Ser Asp Ile Arg Glu Met Gln Glu Ala Arg Thr Leu Ala
 325 330 335
 25 Arg Ala Arg Ala His Ala Arg Asp Gln Ala Arg Glu Gln Asp Arg Ala
 340 345 350
 His Ala Cys Pro Ala Val Glu Glu Thr Pro Met Asn Val Arg Asn Val
 355 360 365
 30 Pro Leu Pro Gly Asp Ala Ala Ala Gly His Pro Asp Arg Ala Ser Gly
 370 375 380
 His Pro Lys Pro His Ser Arg Ser Ser Ser Ala Tyr Arg Lys Ser Ala
 385 390 395 400
 Ser Thr His His Lys Ser Val Phe Ser His Ser Lys Ala Ala Ser Gly
 405 410 415
 35 His Leu Lys Pro Val Ser Gly His Ser Lys Pro Ala Ser Gly His Pro
 420 425 430
 Lys Ser Ala Thr Val Tyr Pro Lys Pro Ala Ser Val His Phe Lys Gly

80

	435	440	445
	Asp Ser Val His Phe Lys Gly Asp Ser Val His Phe Lys Pro Asp Ser		
	450	455	460
5	Val His Phe Lys Pro Ala Ser Ser Asn Pro Lys Pro Ile Thr Gly His		
	465	470	475 480
	His Val Ser Ala Gly Ser His Ser Lys Ser Ala Phe Ser Ala Ala Thr		
		485 490	495
	Ser His Pro Lys Pro Ile Lys Pro Ala Thr Ser His Ala Glu Pro Thr		
		500 505	510
10	Thr Ala Asp Tyr Pro Lys Pro Ala Thr Thr Ser His Pro Lys Pro Ala		
		515 520	525
	Ala Ala Asp Asn Pro Glu Leu Ser Ala Ser His Cys Pro Glu Ile Pro		
		530 535	540
15	Ala Ile Ala His Pro Val Ser Asp Asp Ser Asp Leu Pro Glu Ser Ala		
		545 550	555 560
	Ser Ser Pro Ala Ala Gly Pro Thr Lys Pro Ala Ala Ser Gln Leu Glu		
		565 570	575
	Ser Asp Thr Ile Ala Asp Leu Pro Asp Pro Thr Val Val Thr Thr Ser		
		580 585	590
20	Thr Asn Asp Tyr His Asp Val Val Val Val Asp Val Glu Asp Asp Pro		
		595 600	605
	Asp Glu Met Ala Val		
	610		

(102) INFORMATION FOR SEQ ID NO:101:

- 25 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 32 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- 30 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

TCCAAGCTTC GCCATGGGAC ATAACGGGAG CT

32

(103) INFORMATION FOR SEQ ID NO:102:

- 35 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CGTGAATTCC AAGAATTAC AATCCTTGCT

30

5 (104) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1548 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ATGGGACATA ACGGGAGCTG GATCTCTCCA AATGCCAGCG AGCCGCACAA CGCGTCCGGC 60
GCCGAGGCTG CGGGTGTGAA CCGCAGCGCG CTCGGGGAGT TCGGCGAGGC GCAGCTGTAC 120
15 CGCCAGTTCA CCACCACCGT GCAGGTCGTC ATCTTCATAG GCTCGCTGCT CGGAAACTTC 180
ATGGTGTTAT GGTCAACTTG CCGCACAACC GTGTTCAAAT CTGTCACCAA CAGGTTCAAT 240
AAAAACCTGG CTGCTCGGG GATTGTGCC AGCCTGGTCT GTGTGCCCTT CGACATCATC 300
CTCAGACCA GTCTCACTG TTGCTGGTGG ATCTACACCA TGCTCTTCTG CAAGGTCGTC 360
AAATTTTTC ACAAAGTATT CTGCTCTGTG ACCATCCTCA GCTTCCCTGC TATTGCTTTG 420
20 GACAGGTACT ACTCAGTCCT CTATCCACTG GAGAGGAAAA TATCTGATGC CAAGTCCCGT 480
GAACTGGTGA TGTACATCTG GGCCCATGCA GTGGTGGCCA GTGTCCCTGT GTTTCAGTA 540
ACCAATGTGG CTGACATCTA TGCCACGTCC ACCTGCACGG AAGTCTGGAG CAACTCCTTG 600
GGCCACCTGG TGTACGTTCT GGTGTATAAC ATCACCACGG TCATTGTGCC TGTGGTGGTG 660
GTGTTCTCT TCTTGATACT GATCCGACGG GCCCTGAGTG CCAGCCAGAA GAAGAAGGTC 720
25 ATCATAGCAG CGCTCCGGAC CCCACAGAAC ACCATCTCTA TTCCCTATGC CTCCCAGCGG 780
GAGGCCGAGC TGCACGCCAC CCTGCTCTCC ATGGTGATGG TCTTCATCTT GTGTAGCGTG 840
CCCTATGCCA CCCTGGTCGT CTACCAGACT GTGCTCAATG TCCCTGACAC TTCCGTCTTC 900
TTGCTGCTCA CTGCTGTTTG GCTGCCCAA GTCTCCCTGC TGGCAAACCC TGTTCTCTTT 960
CTTACTGTGA ACAAATCTGT CCGCAAGTGC TTGATAGGGA CCCTGGTGCA ACTACACCAC 1020
30 CGGTACAGTC GCCGTAATGT GGTCAGTACA GGGAGTGGCA TGGCTGAGGC CAGCCTGGAA 1080

CCCAGCATAC GCTCGGGTAG CCAGCTCCTG GAGATGTTCC ACATTGGGCA GCAGCAGATC 1140
 TTTAAGCCCA CAGAGGATGA GGAAGAGAGT GAGGCCAAGT ACATTGGCTC AGCTGACTTC 1200
 CAGGCCAAGG AGATATTTAG CACCTGCCTC GAGGGAGAGC AGGGGCCACA GTTTGCGCCC 1260
 TCTGCCCCAC CCCTGAGCAC AGTGGACTCT GTATCCCAGG TGGCACCAGC AGCCCCTGTG 1320
 5 GAACCTGAAA CATTCCCTGA TAAGTATTCC CTGCAGTTTG GCTTTGGGCC TTTTGAGTTG 1380
 CCTCCTCAGT GGCTCTCAGA GACCCGAAAC AGCAAGAAGC GGCTGCTTCC CCCCTTGGGC 1440
 AACACCCCAG AAGAGCTGAT CCAGACAAAG GTGCCCAAGG TAGGCAGGGT GGAGCGGAAG 1500
 ATGAGCAGAA ACAATAAAGT GAGCATTTTT CCAAAGGTGG ATTCCTAG 1548

(105) INFORMATION FOR SEQ ID NO:104:

- 10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 515 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant
- 15 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:
- | | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| Met | Gly | His | Asn | Gly | Ser | Trp | Ile | Ser | Pro | Asn | Ala | Ser | Glu | Pro | His | 1 | 5 | 10 | 15 |
| Asn | Ala | Ser | Gly | Ala | Glu | Ala | Ala | Gly | Val | Asn | Arg | Ser | Ala | Leu | Gly | 20 | 25 | 30 | |
| Glu | Phe | Gly | Glu | Ala | Gln | Leu | Tyr | Arg | Gln | Phe | Thr | Thr | Thr | Val | Gln | 35 | 40 | 45 | |
| Val | Val | Ile | Phe | Ile | Gly | Ser | Leu | Leu | Gly | Asn | Phe | Met | Val | Leu | Trp | 50 | 55 | 60 | |
| Ser | Thr | Cys | Arg | Thr | Thr | Val | Phe | Lys | Ser | Val | Thr | Asn | Arg | Phe | Ile | 65 | 70 | 75 | 80 |
| Lys | Asn | Leu | Ala | Cys | Ser | Gly | Ile | Cys | Ala | Ser | Leu | Val | Cys | Val | Pro | 85 | 90 | 95 | |
| Phe | Asp | Ile | Ile | Leu | Ser | Thr | Ser | Pro | His | Cys | Cys | Trp | Trp | Ile | Tyr | 100 | 105 | 110 | |
| Thr | Met | Leu | Phe | Cys | Lys | Val | Val | Lys | Phe | Leu | His | Lys | Val | Phe | Cys | 115 | 120 | 125 | |
| Ser | Val | Thr | Ile | Leu | Ser | Phe | Pro | Ala | Ile | Ala | Leu | Asp | Arg | Tyr | Tyr | 130 | 135 | 140 | |

83

	Ser Val Leu Tyr Pro Leu Glu Arg Lys Ile Ser Asp Ala Lys Ser Arg	
	145	150 155 160
	Glu Leu Val Met Tyr Ile Trp Ala His Ala Val Val Ala Ser Val Pro	
		165 170 175
5	Val Phe Ala Val Thr Asn Val Ala Asp Ile Tyr Ala Thr Ser Thr Cys	
		180 185 190
	Thr Glu Val Trp Ser Asn Ser Leu Gly His Leu Val Tyr Val Leu Val	
		195 200 205
10	Tyr Asn Ile Thr Thr Val Ile Val Pro Val Val Val Phe Leu Phe	
		210 215 220
	Leu Ile Leu Ile Arg Arg Ala Leu Ser Ala Ser Gln Lys Lys Lys Val	
		225 230 235 240
	Ile Ile Ala Ala Leu Arg Thr Pro Gln Asn Thr Ile Ser Ile Pro Tyr	
		245 250 255
15	Ala Ser Gln Arg Glu Ala Glu Leu His Ala Thr Leu Leu Ser Met Val	
		260 265 270
	Met Val Phe Ile Leu Cys Ser Val Pro Tyr Ala Thr Leu Val Val Tyr	
		275 280 285
20	Gln Thr Val Leu Asn Val Pro Asp Thr Ser Val Phe Leu Leu Leu Thr	
		290 295 300
	Ala Val Trp Leu Pro Lys Val Ser Leu Leu Ala Asn Pro Val Leu Phe	
		305 310 315 320
	Leu Thr Val Asn Lys Ser Val Arg Lys Cys Leu Ile Gly Thr Leu Val	
		325 330 335
25	Gln Leu His His Arg Tyr Ser Arg Arg Asn Val Val Ser Thr Gly Ser	
		340 345 350
	Gly Met Ala Glu Ala Ser Leu Glu Pro Ser Ile Arg Ser Gly Ser Gln	
		355 360 365
30	Leu Leu Glu Met Phe His Ile Gly Gln Gln Gln Ile Phe Lys Pro Thr	
		370 375 380
	Glu Asp Glu Glu Glu Ser Glu Ala Lys Tyr Ile Gly Ser Ala Asp Phe	
		385 390 395 400
	Gln Ala Lys Glu Ile Phe Ser Thr Cys Leu Glu Gly Glu Gln Gly Pro	
		405 410 415
35	Gln Phe Ala Pro Ser Ala Pro Pro Leu Ser Thr Val Asp Ser Val Ser	
		420 425 430
	Gln Val Ala Pro Ala Ala Pro Val Glu Pro Glu Thr Phe Pro Asp Lys	

84

435 440 445

Tyr Ser Leu Gln Phe Gly Phe Gly Pro Phe Glu Leu Pro Pro Gln Trp
450 455 460

5 Leu Ser Glu Thr Arg Asn Ser Lys Lys Arg Leu Leu Pro Pro Leu Gly
465 470 475 480

Asn Thr Pro Glu Glu Leu Ile Gln Thr Lys Val Pro Lys Val Gly Arg
485 490 495

Val Glu Arg Lys Met Ser Arg Asn Asn Lys Val Ser Ile Phe Pro Lys
500 505 510

10 Val Asp Ser
515

(106) INFORMATION FOR SEQ ID NO:105:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

20 GGAGAATTCA CTAGGCGAGG CGCTCCATC 29

(107) INFORMATION FOR SEQ ID NO:106:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GGAGGATCCA GGAAACCTTA GGCCGAGTCC 30

30 (108) INFORMATION FOR SEQ ID NO:107:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1164 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

ATGAATCGGC ACCATCTGCA GGATCACTTT CTGGAAATAG ACAAGAAGAA CTGCTGTGTG 60
TTCCGAGATG ACTTCATTGC CAAGGTGTTG CCGCCGGTGT TGGGGCTGGA GTTTATCTTT 120
GGGCTTCTGG GCAATGGCCT TGCCCTGTGG ATTTTCTGTT TCCACCTCAA GTCCTGGAAA 180
5 TCCAGCCGGA TTTTCCTGTT CAACCTGGCA GTAGCTGACT TTCTACTGAT CATCTGCCTG 240
CCGTTCTGTA TGGACTACTA TGTGCGGCGT TCAGACTGGA ACTTTGGGGA CATCCCTTGC 300
CGGCTGGTGC TCTTCATGTT TGCCATGAAC CGCCAGGGCA GCATCATCTT CCTCACGGTG 360
GTGGCGGTAG ACAGGTATTT CCGGGTGGTC CATCCCCACC ACGCCCTGAA CAAGATCTCC 420
AATTGGACAG CAGCCATCAT CTCTGCCTT CTGTGGGGCA TCACTGTTGG CCTAACAGTC 480
10 CACCTCCTGA AGAAGAAGTT GCTGATCCAG AATGGCCCTG CAAATGTGTG CATCAGCTTC 540
AGCATCTGCC ATACCTTCCG GTGGCAGGAA GCTATGTTCC TCCTGGAGTT CCTCTGCCC 600
CTGGGCATCA TCCTGTTCTG CTCAGCCAGA ATTATCTGGA GCCTGCGGCA GAGACAAATG 660
GACCGGCATG CCAAGATCAA GAGAGCCATC ACCTTCATCA TGGTGGTGGC CATCGTCTTT 720
GTCATCTGCT TCCTTCCCAG CGTGGTTGTG CGGATCCGCA TCTTCTGGCT CCTGCACACT 780
15 TCGGGCACGC AGAATTGTGA AGTGTACCGC TCGGTGGACC TGGCGTTCTT TATCACTCTC 840
AGCTTCACCT ACATGAACAG CATGCTGGAC CCCGTGGTGT ACTACTTCTC CAGCCCATCC 900
TTTCCCAACT TCTTCTCCAC TTTGATCAAC CGCTGCCTCC AGAGGAAGAT GACAGGTGAG 960
CCAGATAATA ACCGCAGCAC GAGCGTCGAG CTCACAGGGG ACCCCAACAA AACCAGAGGC 1020
GCTCCAGAGG CGTTAATGGC CAACTCCGGT GAGCCATGGA GCCCCTCTTA TCTGGGCCCA 1080
20 ACCTCAAATA ACCATTCCAA GAAGGGACAT TGTACCAAG AACCAGCATC TCTGGAGAAA 1140
CAGTTGGGCT GTTGCATCGA GTAA 1164

(109) INFORMATION FOR SEQ ID NO:108:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 387 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

30 Met Asn Arg His His Leu Gln Asp His Phe Leu Glu Ile Asp Lys Lys

87

Phe Ser Thr Leu Ile Asn Arg Cys Leu Gln Arg Lys Met Thr Gly Glu
 305 310 315 320
 Pro Asp Asn Asn Arg Ser Thr Ser Val Glu Leu Thr Gly Asp Pro Asn
 325 330 335
 5 Lys Thr Arg Gly Ala Pro Glu Ala Leu Met Ala Asn Ser Gly Glu Pro
 340 345 350
 Trp Ser Pro Ser Tyr Leu Gly Pro Thr Ser Asn Asn His Ser Lys Lys
 355 360 365
 10 Gly His Cys His Gln Glu Pro Ala Ser Leu Glu Lys Gln Leu Gly Cys
 370 375 380
 Cys Ile Glu
 385

(110) INFORMATION FOR SEQ ID NO:109:

- 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 20 (iv) ANTI-SENSE: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

ACCATGGCTT GCAATGGCAG TGCGGCCAGG GGGCACT

37

(111) INFORMATION FOR SEQ ID NO:110:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 30 (iv) ANTI-SENSE: YES
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

CGACCAGGAC AAACAGCATC TTGGTCACTT GTCTCCGGC

39

(112) INFORMATION FOR SEQ ID NO:111:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GACCAAGATG CTGTTTGTCC TGGTCGTGGT GTTTGGCAT

39

(113) INFORMATION FOR SEQ ID NO:112:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

15 (iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CGGAATTCAG GATGGATCGG TCTCTTGCTG CGCCT

35

(114) INFORMATION FOR SEQ ID NO:113:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

ATGGCTTGCA ATGGCAGTGC GGCCAGGGGG CACTTTGACC CTGAGGACTT GAACCTGACT 60

GACGAGGCAC TGAGACTCAA GTACCTGGGG CCCCAGCAGA CAGAGCTGTT CATGCCCATC 120

TGTGCCACAT ACCTGCTGAT CTTCTGTTGG GGCCTGTGG GCAATGGGCT GACCTGTCTG 180

GTTCATCCTGC GCCACAAGGC CATGCGCAGC CCTACCAACT ACTACCTCTT CAGCCTGGCC 240

30 GTGTCGGACC TGCTGGTGCT GCTGGTGGGC CTGCCCTGG AGCTCTATGA GATGTGGCAC 300

AACTACCCCT TCCTGCTGGG CGTTGGTGGC TGCTATTTCC GCACGCTACT GTTTGAGATG 360

GTCTGCCTGG CCTCAGTGCT CAACGTCACT GCCCTGAGCG TGAACGCTA TGTGGCCGTG 420

GTGCACCCAC TCCAGGCCAG GTCCATGGTG ACGCGGGCCC ATGTGCGCCG AGTGCTTGGG 480

```

GCCGTCTGGG GTCTTGCCAT GCTCTGCTCC CTGCCCAACA CCAGCCTGCA CGGCATCCGG 540
CAGCTGCACG TGCCCTGCCG GGGCCCAGTG CCAGACTCAG CTGTTTGCAT GCTGGTCCGC 600
CCACGGGCCC TCTACAACAT GGTAGTGAG ACCACCGCGC TGCTCTTCTT CTGCCTGCCC 660
ATGGCCATCA TGAGCGTGCT CTACCTGCTC ATTGGGCTGC GACTGCGGCG GGAGAGGCTG 720
5 CTGCTCATGC AGGAGGCCAA GGGCAGGGGC TCTGCAGCAG CCAGGTCCAG ATACACCTGC 780
AGGCTCCAGC AGCAGGATCG GGGCCGGAGA CAAGTGACCA AGATGCTGTT TGTCTGGTTC 840
GTGGTGTTTG GCATCTGCTG GGCCCCGTTT CACGCCGACC GCGTCATGTG GAGCGTCGTG 900
TCACAGTGGA CAGATGGCCT GCACCTGGCC TTCCAGCACG TGCACGTCAT CTCCGGCATC 960
TTCTTCTACC TGGGCTCGGC GGCCAACCCC GTGCTCTATA GCCTCATGTC CAGCCGCTTC 1020
10 CGAGAGACCT TCCAGGAGGC CCTGTGCCTC GGGGCCTGCT GCCATCGCCT CAGACCCCGC 1080
CACAGCTCCC ACAGCCTCAG CAGGATGACC ACAGGCAGCA CCCTGTGTGA TGTGGGCTCC 1140
CTGGGCAGCT GGGTCCACCC CCTGGCTGGG AACGATGGCC CAGAGGCGCA GCAAGAGACC 1200
GATCCATCCT GA 1212

```

(115) INFORMATION FOR SEQ ID NO:114:

```

15 (i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 403 amino acids
    (B) TYPE: amino acid
    (C) STRANDEDNESS:
    (D) TOPOLOGY: not relevant

20 (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Met Ala Cys Asn Gly Ser Ala Ala Arg Gly His Phe Asp Pro Glu Asp
1          5          10          15
Leu Asn Leu Thr Asp Glu Ala Leu Arg Leu Lys Tyr Leu Gly Pro Gln
25          20          25          30
Gln Thr Glu Leu Phe Met Pro Ile Cys Ala Thr Tyr Leu Leu Ile Phe
35          40          45
Val Val Gly Ala Val Gly Asn Gly Leu Thr Cys Leu Val Ile Leu Arg
50          55          60
30 His Lys Ala Met Arg Thr Pro Thr Asn Tyr Tyr Leu Phe Ser Leu Ala
65          70          75          80
Val Ser Asp Leu Leu Val Leu Leu Val Gly Leu Pro Leu Glu Leu Tyr
85          90          95

```

90

Glu Met Trp His Asn Tyr Pro Phe Leu Leu Gly Val Gly Gly Cys Tyr
 100 105 110
 Phe Arg Thr Leu Leu Phe Glu Met Val Cys Leu Ala Ser Val Leu Asn
 115 120 125
 5 Val Thr Ala Leu Ser Val Glu Arg Tyr Val Ala Val Val His Pro Leu
 130 135 140
 Gln Ala Arg Ser Met Val Thr Arg Ala His Val Arg Arg Val Leu Gly
 145 150 155 160
 10 Ala Val Trp Gly Leu Ala Met Leu Cys Ser Leu Pro Asn Thr Ser Leu
 165 170 175
 His Gly Ile Arg Gln Leu His Val Pro Cys Arg Gly Pro Val Pro Asp
 180 185 190
 Ser Ala Val Cys Met Leu Val Arg Pro Arg Ala Leu Tyr Asn Met Val
 195 200 205
 15 Val Gln Thr Thr Ala Leu Leu Phe Phe Cys Leu Pro Met Ala Ile Met
 210 215 220
 Ser Val Leu Tyr Leu Leu Ile Gly Leu Arg Leu Arg Arg Glu Arg Leu
 225 230 235 240
 20 Leu Leu Met Gln Glu Ala Lys Gly Arg Gly Ser Ala Ala Ala Arg Ser
 245 250 255
 Arg Tyr Thr Cys Arg Leu Gln Gln His Asp Arg Gly Arg Arg Gln Val
 260 265 270
 Thr Lys Met Leu Phe Val Leu Val Val Val Phe Gly Ile Cys Trp Ala
 275 280 285
 25 Pro Phe His Ala Asp Arg Val Met Trp Ser Val Val Ser Gln Trp Thr
 290 295 300
 Asp Gly Leu His Leu Ala Phe Gln His Val His Val Ile Ser Gly Ile
 305 310 315 320
 30 Phe Phe Tyr Leu Gly Ser Ala Ala Asn Pro Val Leu Tyr Ser Leu Met
 325 330 335
 Ser Ser Arg Phe Arg Glu Thr Phe Gln Glu Ala Leu Cys Leu Gly Ala
 340 345 350
 Cys Cys His Arg Leu Arg Pro Arg His Ser Ser His Ser Leu Ser Arg
 355 360 365
 35 Met Thr Thr Gly Ser Thr Leu Cys Asp Val Gly Ser Leu Gly Ser Trp
 370 375 380
 Val His Pro Leu Ala Gly Asn Asp Gly Pro Glu Ala Gln Gln Glu Thr

385

390

395

400

Asp Pro Ser

(116) INFORMATION FOR SEQ ID NO:115:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

GGAAGCTTCA GGCCCAAAGA TGGGGAACAT 30

(117) INFORMATION FOR SEQ ID NO:116:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GTGGATCCAC CCGCGGAGGA CCCAGGCTAG 30

(118) INFORMATION FOR SEQ ID NO:117:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1098 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

30 ATGGGGAACA TCACTGCAGA CAACTCCTCG ATGAGCTGTA CCATCGACCA TACCATCCAC 60
CAGACGCTGG CCCCGGTGGT CTATGTTACC GTGCTGGTGG TGGGCTTCCC GGCCAACTGC 120
CTGTCCCTCT ACTTCGGCTA CCTGCAGATC AAGGCCCGGA ACGAGCTGGG CGTGTACCTG 180
TGCAACCTGA CGGTGGCCGA CCTCTTCTAC ATCTGCTCGC TGCCCTTCTG GCTGCAGTAC 240
GTGCTGCAGC ACGACAACTG GTCTCACGGC GACCTGTCCT GCCAGGTGTG CGGCATCCTC 300
35 CTGTACGAGA ACATCTACAT CAGCGTGGGC TTCCTCTGCT GCATCTCCGT GGACCGCTAC 360

CTGGCTGTGG CCCATCCCTT CCGCTTCCAC CAGTTCCGGA CCCTGAAGGC GGCCGTCGGC 420
 GTCAGCGTGG TCATCTGGGC CAAGGAGCTG CTGACCAGCA TCTACTTCCT GATGCACGAG 480
 GAGGTCATCG AGGACGAGAA CCAGCACCGC GTGTGCTTTG AGCACTACCC CATCCAGGCA 540
 TGGCAGCGCG CCATCAACTA CTACCGCTTC CTGGTGGGCT TCCTCTTCCC CATCTGCCTG 600
 5 CTGCTGGCGT CCTACCAGGG CATCCTGCGC GCCGTGCGCC GGAGCCACGG CACCCAGAAG 660
 AGCCGCAAGG ACCAGATCCA GCGGCTGGTG CTCAGCACCG TGGTCATCTT CCTGGCCTGC 720
 TTCCTGCCCT ACCACGTGTT GCTGCTGGTG CGCAGCGTCT GGGAGGCCAG CTGCGACTTC 780
 GCCAAGGGCG TTTCAACGC CTACCACTTC TCCCTCCTGC TCACCAGCTT CAACTGCGTC 840
 GCCGACCCCG TGCTCTACTG CTTCGTCAGC GAGACCACCC ACCGGGACCT GGCCCGCCTC 900
 10 CGCGGGGCCT GCCTGGCCTT CCTCACCTGC TCCAGGACCG GCCGGGCCAG GGAGGCCTAC 960
 CCGCTGGGTG CCCCCGAGGC CTCCGGGAAA AGCGGGGCC AGGGTGAGGA GCCCCAGCTG 1020
 TTGACCAAGC TCCACCCGGC CTTCAGACC CTTAACTCGC CAGGGTCGGG CGGGTTCCCC 1080
 ACGGGCAGGT TGGCCTAG 1098

(119) INFORMATION FOR SEQ ID NO:118:

- 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 365 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant
- 20 (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:
- | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Asn | Ile | Thr | Ala | Asp | Asn | Ser | Ser | Met | Ser | Cys | Thr | Ile | Asp |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| His | Thr | Ile | His | Gln | Thr | Leu | Ala | Pro | Val | Val | Tyr | Val | Thr | Val | Leu |
| 25 | | | 20 | | | | | 25 | | | | | 30 | | |
| Val | Val | Gly | Phe | Pro | Ala | Asn | Cys | Leu | Ser | Leu | Tyr | Phe | Gly | Tyr | Leu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Gln | Ile | Lys | Ala | Arg | Asn | Glu | Leu | Gly | Val | Tyr | Leu | Cys | Asn | Leu | Thr |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Val | Ala | Asp | Leu | Phe | Tyr | Ile | Cys | Ser | Leu | Pro | Phe | Trp | Leu | Gln | Tyr |
| 30 | 65 | | | | 70 | | | | | 75 | | | | 80 | |
| Val | Leu | Gln | His | Asp | Asn | Trp | Ser | His | Gly | Asp | Leu | Ser | Cys | Gln | Val |
| | | | 85 | | | | | | 90 | | | | | 95 | |

	Cys Gly Ile Leu Leu Tyr Glu Asn Ile Tyr Ile Ser Val Gly Phe Leu	
	100	105 110
	Cys Cys Ile Ser Val Asp Arg Tyr Leu Ala Val Ala His Pro Phe Arg	
	115	120 125
5	Phe His Gln Phe Arg Thr Leu Lys Ala Ala Val Gly Val Ser Val Val	
	130	135 140
	Ile Trp Ala Lys Glu Leu Leu Thr Ser Ile Tyr Phe Leu Met His Glu	
	145	150 155 160
10	Glu Val Ile Glu Asp Glu Asn Gln His Arg Val Cys Phe Glu His Tyr	
	165	170 175
	Pro Ile Gln Ala Trp Gln Arg Ala Ile Asn Tyr Tyr Arg Phe Leu Val	
	180	185 190
	Gly Phe Leu Phe Pro Ile Cys Leu Leu Leu Ala Ser Tyr Gln Gly Ile	
	195	200 205
15	Leu Arg Ala Val Arg Arg Ser His Gly Thr Gln Lys Ser Arg Lys Asp	
	210	215 220
	Gln Ile Gln Arg Leu Val Leu Ser Thr Val Val Ile Phe Leu Ala Cys	
	225	230 235 240
20	Phe Leu Pro Tyr His Val Leu Leu Leu Val Arg Ser Val Trp Glu Ala	
	245	250 255
	Ser Cys Asp Phe Ala Lys Gly Val Phe Asn Ala Tyr His Phe Ser Leu	
	260	265 270
	Leu Leu Thr Ser Phe Asn Cys Val Ala Asp Pro Val Leu Tyr Cys Phe	
	275	280 285
25	Val Ser Glu Thr Thr His Arg Asp Leu Ala Arg Leu Arg Gly Ala Cys	
	290	295 300
	Leu Ala Phe Leu Thr Cys Ser Arg Thr Gly Arg Ala Arg Glu Ala Tyr	
	305	310 315 320
30	Pro Leu Gly Ala Pro Glu Ala Ser Gly Lys Ser Gly Ala Gln Gly Glu	
	325	330 335
	Glu Pro Glu Leu Leu Thr Lys Leu His Pro Ala Phe Gln Thr Pro Asn	
	340	345 350
	Ser Pro Gly Ser Gly Gly Phe Pro Thr Gly Arg Leu Ala	
	355	360 365

35 (120) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

GACCTCGAGT CCTTCTACAC CTCATC

26

(121) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

15 TGCTCTAGAT TCCAGATAGG TGAAACTTG

30

(122) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

20 (A) LENGTH: 1416 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

ATGGATATTC TTTGTGAAGA AAATACTTCT TTGAGCTCAA CTACGAACTC CCTAATGCAA 60

25 TTAAATGATG ACAACAGGCT CTACAGTAAT GACTTTAACT CCGGAGAAGC TAACACTTCT 120

GATGCATTTA ACTGGACAGT CGACTCTGAA AATCGAACCA ACCTTTCCTG TGAAGGGTGC 180

CTCTCACCGT CGTGTCTCTC CTTACTTCAT CTCCAGGAAA AAAACTGGTC TGCTTTACTG 240

ACAGCCGTAG TGATTATTCT AACTATTGCT GGAAACATAC TCGTCATCAT GGCAGTGTCC 300

CTAGAGAAAA AGCTGCAGAA TGCCACCAAC TATTTCTGA TGCACTTGC CATAGCTGAT 360

30 ATGCTGCTGG GTTTCCTTGT CATGCCCGTG TCCATGTTAA CCATCCTGTA TGGGTACCGG 420

TGGCCTCTGC CGAGCAAGCT TTGTGCAGTC TGGATTTACC TGGACGTGCT CTTCTCCACG 480

GCCTCCATCA TGCACCTCTG CGCCATCTCG CTGGACCGCT ACGTCGCCAT CCAGAATCCC 540

ATCCACCACA GCCGCTTCAA CTCCAGAACT AAGGCATTTT TGAAAATCAT TGCTGTTTGG 600

ACCATATCAG TAGGTATATC CATGCCAATA CCAGTCTTTG GGCTACAGGA CGATTCTGAAG 660
 GTCTTTAAGG AGGGGAGTTG CTTACTCGCC GATGATAACT TTGTCCTGAT CGGCTCTTTT 720
 GTGTCATTTT TCATTCCCTT AACCATCATG GTGATCACCT ACTTTCTAAC TATCAAGTCA 780
 CTCCAGAAAG AAGCTACTTT GTGTGTAAGT GATCTTGGCA CACGGGCCAA ATTAGCTTCT 840
 5 TTCAGCTTCC TCCCTCAGAG TTCTTTGTCT TCAGAAAAGC TCTTCCAGCG GTCGATCCAT 900
 AGGGAGCCAG GGTCTACAC AGGCAGGAGG ACTATGCAGT CCATCAGCAA TGAGCAAAAG 960
 GCATGCAAGG TGCTGGGCAT CGTCTTCTTC CTGTTTGTGG TGATGTGGTG CCCTTTCTTC 1020
 ATCACAAACA TCATGGCCGT CATCTGCAAA GAGTCCTGCA ATGAGGATGT CATTGGGGCC 1080
 CTGCTCAATG TGTTTGTTTG GATCGGTTAT CTCTCTCAG CAGTCAACCC ACTAGTCTAC 1140
 10 ACACTGTTCA ACAAGACCTA TAGGTCAGCC TTTTCACGGT ATATTCAGTG TCAGTACAAG 1200
 GAAAACAAAA AACCATTGCA GTTAATTTTA GTGAACACAA TACCGGCTTT GGCCTACAAG 1260
 TCTAGCCAAC TTCAAATGGG ACAAAAAAAG AATTCAAAGC AAGATGCCAA GACAACAGAT 1320
 AATGACTGCT CAATGGTTGC TCTAGGAAAG CAGTATTCTG AAGAGGCTTC TAAAGACAAT 1380
 AGCGACGGAG TGAATGAAAA GGTGAGCTGT GTGTGA 1416

15 (123) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 471 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

20 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Met Asp Ile Leu Cys Glu Glu Asn Thr Ser Leu Ser Ser Thr Thr Asn
 1 5 10 15
 25 Ser Leu Met Gln Leu Asn Asp Asp Asn Arg Leu Tyr Ser Asn Asp Phe
 20 25 30
 Asn Ser Gly Glu Ala Asn Thr Ser Asp Ala Phe Asn Trp Thr Val Asp
 35 40 45
 30 Ser Glu Asn Arg Thr Asn Leu Ser Cys Glu Gly Cys Leu Ser Pro Ser
 50 55 60
 Cys Leu Ser Leu Leu His Leu Gln Glu Lys Asn Trp Ser Ala Leu Leu
 65 70 75 80

Thr Ala Val Val Ile Ile Leu Thr Ile Ala Gly Asn Ile Leu Val Ile
 85 90 95
 Met Ala Val Ser Leu Glu Lys Lys Leu Gln Asn Ala Thr Asn Tyr Phe
 100 105 110
 5 Leu Met Ser Leu Ala Ile Ala Asp Met Leu Leu Gly Phe Leu Val Met
 115 120 125
 Pro Val Ser Met Leu Thr Ile Leu Tyr Gly Tyr Arg Trp Pro Leu Pro
 130 135 140
 10 Ser Lys Leu Cys Ala Val Trp Ile Tyr Leu Asp Val Leu Phe Ser Thr
 145 150 155 160
 Ala Ser Ile Met His Leu Cys Ala Ile Ser Leu Asp Arg Tyr Val Ala
 165 170 175
 Ile Gln Asn Pro Ile His His Ser Arg Phe Asn Ser Arg Thr Lys Ala
 180 185 190
 15 Phe Leu Lys Ile Ile Ala Val Trp Thr Ile Ser Val Gly Ile Ser Met
 195 200 205
 Pro Ile Pro Val Phe Gly Leu Gln Asp Asp Ser Lys Val Phe Lys Glu
 210 215 220
 20 Gly Ser Cys Leu Leu Ala Asp Asp Asn Phe Val Leu Ile Gly Ser Phe
 225 230 235 240
 Val Ser Phe Phe Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Phe Leu
 245 250 255
 Thr Ile Lys Ser Leu Gln Lys Glu Ala Thr Leu Cys Val Ser Asp Leu
 260 265 270
 25 Gly Thr Arg Ala Lys Leu Ala Ser Phe Ser Phe Leu Pro Gln Ser Ser
 275 280 285
 Leu Ser Ser Glu Lys Leu Phe Gln Arg Ser Ile His Arg Glu Pro Gly
 290 295 300
 30 Ser Tyr Thr Gly Arg Arg Thr Met Gln Ser Ile Ser Asn Glu Gln Lys
 305 310 315 320
 Ala Cys Lys Val Leu Gly Ile Val Phe Phe Leu Phe Val Val Met Trp
 325 330 335
 Cys Pro Phe Phe Ile Thr Asn Ile Met Ala Val Ile Cys Lys Glu Ser
 340 345 350
 35 Cys Asn Glu Asp Val Ile Gly Ala Leu Leu Asn Val Phe Val Trp Ile
 355 360 365
 Gly Tyr Leu Ser Ser Ala Val Asn Pro Leu Val Tyr Thr Leu Phe Asn

97

	370		375		380
	Lys Thr Tyr Arg Ser Ala Phe Ser Arg Tyr Ile Gln Cys Gln Tyr Lys				
	385		390		395 400
5	Glu Asn Lys Lys Pro Leu Gln Leu Ile Leu Val Asn Thr Ile Pro Ala				
		405		410	415
	Leu Ala Tyr Lys Ser Ser Gln Leu Gln Met Gly Gln Lys Lys Asn Ser				
		420		425	430
	Lys Gln Asp Ala Lys Thr Thr Asp Asn Asp Cys Ser Met Val Ala Leu				
		435		440	445
10	Gly Lys Gln Tyr Ser Glu Glu Ala Ser Lys Asp Asn Ser Asp Gly Val				
		450		455	460
	Asn Glu Lys Val Ser Cys Val				
		465		470	

(124) INFORMATION FOR SEQ ID NO:123:

- 15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- 20 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

GACCTCGAGG TTGCTTAAGA CTGAAGC

27

(125) INFORMATION FOR SEQ ID NO:124:

- 25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

- 30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

ATTCTAGAC ATATGTAGCT TGTACCG

27

(126) INFORMATION FOR SEQ ID NO:125:

- 35 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1377 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

	ATGGTGAACC	TGAGGAATGC	GGTGCATTCA	TTCCTTGTGC	ACCTAATTGG	CCTATTGGTT	60
	TGGCAATGTG	ATATTTCTGT	GAGCCCAGTA	GCAGCTATAG	TAAGTGACAT	TTTCAATACC	120
5	TCCGATGGTG	GACGCTTCAA	ATTCCCAGAC	GGGGTACAAA	ACTGGCCAGC	ACTTTCAATC	180
	GTCATCATAA	TAATCATGAC	AATAGGTGGC	AACATCCTTG	TGATCATGGC	AGTAAGCATG	240
	GAAAAGAAAC	TGCACAATGC	CACCAATTAC	TTCTTAATGT	CCCTAGCCAT	TGCTGATATG	300
	CTAGTGGGAC	TACTTGTCAT	GCCCCTGTCT	CTCCTGGCAA	TCCTTTATGA	TTATGTCTGG	360
	CCACTACCTA	GATATTTGTG	CCCCGTCTGG	ATTTCTTTAG	ATGTTTTATT	TTCAACAGCG	420
10	TCCATCATGC	ACCTCTGCGC	TATATCGCTG	GATCGGTATG	TAGCAATACG	TAATCCTATT	480
	GAGCATAGCC	GTTTCAATTC	GCGGACTAAG	GCCATCATGA	AGATTGCTAT	TGTTTGGGCA	540
	ATTTCTATAG	GTGTATCAGT	TCCTATCCCT	GTGATTGGAC	TGAGGGACGA	AGAAAAGGTG	600
	TTCGTGAACA	ACACGACGTG	CGTGCTCAAC	GACCCAAATT	TCGTTCTTAT	TGGGTCCTTC	660
	GTAGCTTTCT	TCATACCGCT	GACGATTATG	GTGATTACGT	ATTGCCTGAC	CATCTACGTT	720
15	CTGCGCCGAC	AAGCTTTGAT	GTTACTGCAC	GGCCACACCG	AGGAACCGCC	TGGACTAAGT	780
	CTGGATTTCC	TGAAGTGCTG	CAAGAGGAAT	ACGGCCGAGG	AAGAGAACTC	TGCAAACCCCT	840
	AACCAAGACC	AGAACGCACG	CCGAAGAAAG	AAGAAGGAGA	GACGTCCTAG	GGGCACCATG	900
	CAGGCTATCA	ACAATGAAAG	AAAAGCTTCG	AAAGTCCTTG	GGATTGTTTT	CTTTGTGTTT	960
	CTGATCATGT	GGTGCCCATT	TTTCATTACC	AATATTCTGT	CTGTTCTTTG	TGAGAAGTCC	1020
20	TGTAACCAAA	AGCTCATGGA	AAAGCTTCTG	AATGTGTTTG	TTTGGATTGG	CTATGTTTGT	1080
	TCAGGAATCA	ATCCTCTGGT	GTATACTCTG	TTCAACAAAA	TTTACCGAAG	GGCATTCTCC	1140
	AACTATTTGC	GTTGCAATTA	TAAGGTAGAG	AAAAAGCCTC	CTGTCAGGCA	GATTCCAAGA	1200
	GTTGCCGCCA	CTGCTTTGTC	TGGGAGGGAG	CTTAATGTTA	ACATTTATCG	GCATACCAAT	1260
	GAACCGGTGA	TCGAGAAAGC	CAGTGACAAT	GAGCCCGGTA	TAGAGATGCA	AGTTGAGAAT	1320
25	TTAGAGTTAC	CAGTAAATCC	CTCCAGTGTG	GTTAGCGAAA	GGATTAGCAG	TGTGTGA	1377

(127) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 458 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

5	Met Val Asn Leu Arg Asn Ala Val His Ser Phe Leu Val His Leu Ile	1	5	10	15
	Gly Leu Leu Val Trp Gln Cys Asp Ile Ser Val Ser Pro Val Ala Ala	20	25	30	
10	Ile Val Thr Asp Ile Phe Asn Thr Ser Asp Gly Gly Arg Phe Lys Phe	35	40	45	
	Pro Asp Gly Val Gln Asn Trp Pro Ala Leu Ser Ile Val Ile Ile Ile	50	55	60	
	Ile Met Thr Ile Gly Gly Asn Ile Leu Val Ile Met Ala Val Ser Met	65	70	75	80
15	Glu Lys Lys Leu His Asn Ala Thr Asn Tyr Phe Leu Met Ser Leu Ala	85	90	95	
	Ile Ala Asp Met Leu Val Gly Leu Leu Val Met Pro Leu Ser Leu Leu	100	105	110	
20	Ala Ile Leu Tyr Asp Tyr Val Trp Pro Leu Pro Arg Tyr Leu Cys Pro	115	120	125	
	Val Trp Ile Ser Leu Asp Val Leu Phe Ser Thr Ala Ser Ile Met His	130	135	140	
	Leu Cys Ala Ile Ser Leu Asp Arg Tyr Val Ala Ile Arg Asn Pro Ile	145	150	155	160
25	Glu His Ser Arg Phe Asn Ser Arg Thr Lys Ala Ile Met Lys Ile Ala	165	170	175	
	Ile Val Trp Ala Ile Ser Ile Gly Val Ser Val Pro Ile Pro Val Ile	180	185	190	
30	Gly Leu Arg Asp Glu Glu Lys Val Phe Val Asn Asn Thr Thr Cys Val	195	200	205	
	Leu Asn Asp Pro Asn Phe Val Leu Ile Gly Ser Phe Val Ala Phe Phe	210	215	220	
	Ile Pro Leu Thr Ile Met Val Ile Thr Tyr Cys Leu Thr Ile Tyr Val	225	230	235	240
35	Leu Arg Arg Gln Ala Leu Met Leu Leu His Gly His Thr Glu Glu Pro	245	250	255	

100

Pro Gly Leu Ser Leu Asp Phe Leu Lys Cys Cys Lys Arg Asn Thr Ala
 260 265 270

Glu Glu Glu Asn Ser Ala Asn Pro Asn Gln Asp Gln Asn Ala Arg Arg
 275 280 285

5 Arg Lys Lys Lys Glu Arg Arg Pro Arg Gly Thr Met Gln Ala Ile Asn
 290 295 300

Asn Glu Arg Lys Ala Ser Lys Val Leu Gly Ile Val Phe Phe Val Phe
 305 310 315 320

10 Leu Ile Met Trp Cys Pro Phe Phe Ile Thr Asn Ile Leu Ser Val Leu
 325 330 335

Cys Glu Lys Ser Cys Asn Gln Lys Leu Met Glu Lys Leu Leu Asn Val
 340 345 350

Phe Val Trp Ile Gly Tyr Val Cys Ser Gly Ile Asn Pro Leu Val Tyr
 355 360 365

15 Thr Leu Phe Asn Lys Ile Tyr Arg Arg Ala Phe Ser Asn Tyr Leu Arg
 370 375 380

Cys Asn Tyr Lys Val Glu Lys Lys Pro Pro Val Arg Gln Ile Pro Arg
 385 390 395 400

20 Val Ala Ala Thr Ala Leu Ser Gly Arg Glu Leu Asn Val Asn Ile Tyr
 405 410 415

Arg His Thr Asn Glu Pro Val Ile Glu Lys Ala Ser Asp Asn Glu Pro
 420 425 430

Gly Ile Glu Met Gln Val Glu Asn Leu Glu Leu Pro Val Asn Pro Ser
 435 440 445

25 Ser Val Val Ser Glu Arg Ile Ser Ser Val
 450 455

(128) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

35 GGTAAGCTTG GCAGTCCACG CCAGGCCTTC

30

(129) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

TCCGAATTCT CTGTAGACAC AAGGCTTTGG 30

(130) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

ATGGATCAGT TCCCTGAATC AGTGACAGAA AACTTTGAGT ACGATGATTT GGCTGAGGCC 60

TGTTATATTG GGGACATCGT GGTCTTTGGG ACTGTGTTCC TGTCCATATT CTACTCCGTC 120

ATCTTTGCCA TTGGCCTGGT GGGAAATTG TTGGTAGTGT TTGCCCTCAC CAACAGCAAG 180

AAGCCCCAAGA GTGTCACCGA CATTTACCTC CTGAACCTGG CCTTGTCTGA TCTGCTGTTT 240

GTAGCCACTT TGCCCTTCTG GACTCACTAT TTGATAAATG AAAAGGGCCT CCACAATGCC 300

ATGTGCAAAT TCACTACCGC CTTCTTCTTC ATCGGCTTTT TTGGAAGCAT ATTCTTCATC 360

ACCGTCATCA GCATTGATAG GTACCTGGCC ATCGTCCTGG CCGCCAACTC CATGAACAAC 420

CGGACCGTGC AGCATGGCGT CACCATCAGC CTAGGCGTCT GGGCAGCAGC CATTTTGGTG 480

GCAGCACCCC AGTTCATGTT CACAAAGCAG AAAGAAAATG AATGCCTTGG TGA CTACCCC 540

GAGGTCTCTC AGGAAATCTG GCCCGTGCTC CGCAATGTGG AAACAAATTT TCTTGGCTTC 600

CTACTCCCCC TGCTCATTAT GAGTTATTGC TACTTCAGAA TCATCCAGAC GCTGTTTTCC 660

TGCAAGAACC ACAAGAAAGC CAAAGCCATT AACTGATCC TTCTGGTGGT CATCGTGTTT 720

TTCTCTTCTT GGACACCCTA CAACGTTATG ATTTTCCTGG AGACGCTTAA GCTCTATGAC 780

TTCTTTCCCA GTTGTGACAT GAGGAAGGAT CTGAGGCTGG CCCTCAGTGT GACTGAGACG 840

GTTGCATTTA GCCATTGTTG CCTGAATCCT CTCATCTATG CATTTGCTGG GGAGAAGTTC 900

AGAAGATACC TTTACCACCT GTATGGGAAA TGCCCTGGCTG TCCTGTGTGG GCGCTCAGTC 960

CACGTTGATT TCTCCTCATC TGAATCACAA AGGAGCAGGC ATGGAAGTGT TCTGAGCAGC 1020
 AATTTTACTT ACCACACGAG TGATGGAGAT GCATTGCTCC TTCTCTGA 1068

(131) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:
 5 (A) LENGTH: 355 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Met	Asp	Gln	Phe	Pro	Glu	Ser	Val	Thr	Glu	Asn	Phe	Glu	Tyr	Asp	Asp	1	5	10	15
Leu	Ala	Glu	Ala	Cys	Tyr	Ile	Gly	Asp	Ile	Val	Val	Phe	Gly	Thr	Val	20	25	30	
Phe	Leu	Ser	Ile	Phe	Tyr	Ser	Val	Ile	Phe	Ala	Ile	Gly	Leu	Val	Gly	35	40	45	
Asn	Leu	Leu	Val	Val	Phe	Ala	Leu	Thr	Asn	Ser	Lys	Lys	Pro	Lys	Ser	50	55	60	
Val	Thr	Asp	Ile	Tyr	Leu	Leu	Asn	Leu	Ala	Leu	Ser	Asp	Leu	Leu	Phe	65	70	75	80
Val	Ala	Thr	Leu	Pro	Phe	Trp	Thr	His	Tyr	Leu	Ile	Asn	Glu	Lys	Gly	85	90	95	
Leu	His	Asn	Ala	Met	Cys	Lys	Phe	Thr	Thr	Ala	Phe	Phe	Phe	Ile	Gly	100	105	110	
Phe	Phe	Gly	Ser	Ile	Phe	Phe	Ile	Thr	Val	Ile	Ser	Ile	Asp	Arg	Tyr	115	120	125	
Leu	Ala	Ile	Val	Leu	Ala	Ala	Asn	Ser	Met	Asn	Asn	Arg	Thr	Val	Gln	130	135	140	
His	Gly	Val	Thr	Ile	Ser	Leu	Gly	Val	Trp	Ala	Ala	Ala	Ile	Leu	Val	145	150	155	160
Ala	Ala	Pro	Gln	Phe	Met	Phe	Thr	Lys	Gln	Lys	Glu	Asn	Glu	Cys	Leu	165	170	175	
Gly	Asp	Tyr	Pro	Glu	Val	Leu	Gln	Glu	Ile	Trp	Pro	Val	Leu	Arg	Asn	180	185	190	
Val	Glu	Thr	Asn	Phe	Leu	Gly	Phe	Leu	Leu	Pro	Leu	Leu	Ile	Met	Ser	195	200	205	

103

Tyr Cys Tyr Phe Arg Ile Ile Gln Thr Leu Phe Ser Cys Lys Asn His
210 215 220

Lys Lys Ala Lys Ala Ile Lys Leu Ile Leu Leu Val Val Ile Val Phe
225 230 235 240

5 Phe Leu Phe Trp Thr Pro Tyr Asn Val Met Ile Phe Leu Glu Thr Leu
245 250 255

Lys Leu Tyr Asp Phe Phe Pro Ser Cys Asp Met Arg Lys Asp Leu Arg
260 265 270

10 Leu Ala Leu Ser Val Thr Glu Thr Val Ala Phe Ser His Cys Cys Leu
275 280 285

Asn Pro Leu Ile Tyr Ala Phe Ala Gly Glu Lys Phe Arg Arg Tyr Leu
290 295 300

Tyr His Leu Tyr Gly Lys Cys Leu Ala Val Leu Cys Gly Arg Ser Val
305 310 315 320

15 His Val Asp Phe Ser Ser Ser Glu Ser Gln Arg Ser Arg His Gly Ser
325 330 335

Val Leu Ser Ser Asn Phe Thr Tyr His Thr Ser Asp Gly Asp Ala Leu
340 345 350

20 Leu Leu Leu
355

(132) INFORMATION FOR SEQ ID NO:131:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
25 (C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

GATCTCCAGT AGGCATAAGT GGACAATTCT GG

32

30 (133) INFORMATION FOR SEQ ID NO:132:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
35 (C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

CTCCTTCGGT CCTCCTATCG TTGTCAGAAG

30

(134) INFORMATION FOR SEQ ID NO:133:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

10 AGAAGGCCAA GATCGCGCGG CTGGCCCTCA

30

(135) INFORMATION FOR SEQ ID NO:134:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

CGGCGCCACC GCACGAAAAA GCTCATCTTC

30

20 (136) INFORMATION FOR SEQ ID NO:135:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

GCCAAGAAGC GGGTGAAGTT CCTGGTGGTG GCA

33

(137) INFORMATION FOR SEQ ID NO:136:

- 30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

CAGGCGGAAG GTGAAAGTCC TGGTCCTCGT

30

(138) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

CGGCGCCTGC GGGCCAAGCG GCTGGTGGTG GTG

33

(139) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:

15

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:

20

CCAAGCACAA AGCCAAGAAA GTGACCATCA C

31

(140) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:

25

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

GCGCCGGCGC ACCAAATGCT TGCTGGTGGT

30

30

(141) INFORMATION FOR SEQ ID NO:140:

(i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

106

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:140:

CAAAAAGCTG AAGAAATCTA AGAAGATCAT CTTTATTGTC G

41

(142) INFORMATION FOR SEQ ID NO:141:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:141:

CAAGACCAAG GCAAAACGCA TGATCGCCAT

30

(143) INFORMATION FOR SEQ ID NO:142:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

GTCAAGGAGA AGTCCAAAAG GATCATCATC

30

(144) INFORMATION FOR SEQ ID NO:143:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

30

CGCCGCGTGC GGGCCAAGCA GCTCCTGCTC

30

(145) INFORMATION FOR SEQ ID NO:144:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

107

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:

CCTGATAAGC GCTATAAAAT GGTCTGTTT CGA

33

(146) INFORMATION FOR SEQ ID NO:145:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

GAAAGACAAA AGAGAGTCAA GAGGATGTCT TTATTG

36

(147) INFORMATION FOR SEQ ID NO:146:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

CGGAGAAAGA GGGTGAAACG CACAGCCATC GCC

33

(148) INFORMATION FOR SEQ ID NO:147:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

30 AAGCTTCAGC GGGCCAAGGC ACTGGTCACC

30

(149) INFORMATION FOR SEQ ID NO:148:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

108

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

CAGCGGCAGA AGGC^AAAAG GGTGGCCATC

30

(150) INFORMATION FOR SEQ ID NO:149:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CGGCAGAAGG CGAAGCGCAT GATCCTCGCG

30

(151) INFORMATION FOR SEQ ID NO:150:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GAGCGCAACA AGGCCAAAAA GGTGATCATC

30

(152) INFORMATION FOR SEQ ID NO:151:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:151:

30

GGTGTAACA AAAAGGCTAA AACACAATT ATTCTTATT

39

(153) INFORMATION FOR SEQ ID NO:152:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

109

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:152:

GAGAGCCAGC TCAAGAGCAC CGTGGTG

27

(154) INFORMATION FOR SEQ ID NO:153:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:153:

CCACAAGCAA ACCAAGAAAA TGCTGGCTGT

30

(155) INFORMATION FOR SEQ ID NO:154:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

CATCAAGTGT ATCATGTGCC AAGTACGCCC

30

(156) INFORMATION FOR SEQ ID NO:155:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:155:

30 CTAGAGAGTC AGATGAAGTG TACAGTAGTG GCAC

34

(157) INFORMATION FOR SEQ ID NO:156:

- 35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

110

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

CGGACAAAAG TGAAACTAA AAAGATGTTT CTCATT

36

(158) INFORMATION FOR SEQ ID NO:157:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

10

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GCTGAGGTTT GCAATAAACT AACCATGTTT GTG

33

(159) INFORMATION FOR SEQ ID NO:158:

15

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GGGAGGCCGA GCTGAAAGCC ACCCTGCTC

29

(160) INFORMATION FOR SEQ ID NO:159:

25

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

30 CAAGATCAAG AGAGCCAAAA CCTTCATCAT G

31

(161) INFORMATION FOR SEQ ID NO:160:

35

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

111

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

CCGGAGACAA GTGAAGAGA TGCTGTTTGT C

31

(162) INFORMATION FOR SEQ ID NO:161:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

10 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

GCAAGGACCA GATCAAGCGG CTGGTGCTCA

30

(163) INFORMATION FOR SEQ ID NO:162:

- 15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

CAAGAAAGCC AAAGCCAAGA AACTGATCCT TCTG

34

(164) INFORMATION FOR SEQ ID NO:163:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1068 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

30 ATGGAAGATT TGGAGGAAAC ATTATTTGAA GAATTGAAA ACTATTCCTA TGACCTAGAC 60

TATTACTCTC TGGAGTCTGA TTTGGAGGAG AAAGTCCAGC TGGGAGTTGT TCACTGGGTC 120

TCCCTGGTGT TATATTGTTT GGCTTTTGT CTGGGAATTC CAGGAAATGC CATCGTCATT 180

TGGTTCACGG GGCTCAAGTG GAAGAAGACA GTCACCACTC TGTGGTTCCT CAATCTAGCC 240

ATTGCGGATT TCATTTTCT TCTCTTCTG CCCCTGTACA TCTCCTATGT GGCCATGAAT 300

TTCCACTGGC CCTTTGGCAT CTGGCTGTGC AAAGCCAATT CCTTCACTGC CCAGTTGAAC 360
 ATGTTTGCCA GTGTTTTTTT CCTGACAGTG ATCAGCCTGG ACCACTATAT CCACTTGATC 420
 CATCCTGTCT TATCTCATCG GCATCGAACC CTCAAGAACT CTCTGATTGT CATTATATTC 480
 ATCTGGCTTT TGGCTTCTCT AATTGGCGGT CCTGCCCTGT ACTTCCGGGA CACTGTGGAG 540
 5 TTCAATAATC ATACTCTTTG CTATAACAAT TTTCAGAAGC ATGATCCTGA CCTCACTTTG 600
 ATCAGGCACC ATGTTCTGAC TTGGGTGAAA TTTATCATTG GCTATCTCTT CCCTTTGCTA 660
 ACAATGAGTA TTTGCTACTT GTGTCTCATC TTCAAGGTGA AGAAGCGAAC AGTCCTGATC 720
 TCCAGTAGGC ATAAGTGGAC AATTCTGGTT GTGGTTGTGG CCTTTGTGGT TTGCTGGACT 780
 CCTTATCACC TGTTTAGCAT TTGGGAGCTC ACCATTCAACC ACAATAGCTA TTCCCACCAT 840
 10 GTGATGCAGG CTGGAATCCC CCTCTCCACT GGTTTGGCAT TCCTCAATAG TTGCTTGAAC 900
 CCCATCCTTT ATGTCCTAAT TAGTAAGAAG TTCCAAGCTC GCTTCCGGTC CTCAGTTGCT 960
 GAGATACTCA AGTACACACT GTGGGAAGTC AGCTGTTCTG GCACAGTGAG TGAACAGCTC 1020
 AGGAACTCAG AAACCAAGAA TCTGTGTCTC CTGGAAACAG CTCAATAA 1068

(165) INFORMATION FOR SEQ ID NO:164:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 355 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant
 20 (ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:
 Met Glu Asp Leu Glu Glu Thr Leu Phe Glu Glu Phe Glu Asn Tyr Ser
 1 5 10 15
 Tyr Asp Leu Asp Tyr Tyr Ser Leu Glu Ser Asp Leu Glu Glu Lys Val
 25 20 25 30
 Gln Leu Gly Val Val His Trp Val Ser Leu Val Leu Tyr Cys Leu Ala
 35 40 45
 Phe Val Leu Gly Ile Pro Gly Asn Ala Ile Val Ile Trp Phe Thr Gly
 50 55 60
 30 Leu Lys Trp Lys Lys Thr Val Thr Thr Leu Trp Phe Leu Asn Leu Ala
 65 70 75 80
 Ile Ala Asp Phe Ile Phe Leu Leu Phe Leu Pro Leu Tyr Ile Ser Tyr
 85 90 95

113

Val Ala Met Asn Phe His Trp Pro Phe Gly Ile Trp Leu Cys Lys Ala
100 105 110

Asn Ser Phe Thr Ala Gln Leu Asn Met Phe Ala Ser Val Phe Phe Leu
115 120 125

5 Thr Val Ile Ser Leu Asp His Tyr Ile His Leu Ile His Pro Val Leu
130 135 140

Ser His Arg His Arg Thr Leu Lys Asn Ser Leu Ile Val Ile Ile Phe
145 150 155 160

10 Ile Trp Leu Leu Ala Ser Leu Ile Gly Gly Pro Ala Leu Tyr Phe Arg
165 170 175

Asp Thr Val Glu Phe Asn Asn His Thr Leu Cys Tyr Asn Asn Phe Gln
180 185 190

Lys His Asp Pro Asp Leu Thr Leu Ile Arg His His Val Leu Thr Trp
195 200 205

15 Val Lys Phe Ile Ile Gly Tyr Leu Phe Pro Leu Leu Thr Met Ser Ile
210 215 220

Cys Tyr Leu Cys Leu Ile Phe Lys Val Lys Lys Arg Thr Val Leu Ile
225 230 235 240

20 Ser Ser Arg His Lys Trp Thr Ile Leu Val Val Val Val Ala Phe Val
245 250 255

Val Cys Trp Thr Pro Tyr His Leu Phe Ser Ile Trp Glu Leu Thr Ile
260 265 270

His His Asn Ser Tyr Ser His His Val Met Gln Ala Gly Ile Pro Leu
275 280 285

25 Ser Thr Gly Leu Ala Phe Leu Asn Ser Cys Leu Asn Pro Ile Leu Tyr
290 295 300

Val Leu Ile Ser Lys Lys Phe Gln Ala Arg Phe Arg Ser Ser Val Ala
305 310 315 320

30 Glu Ile Leu Lys Tyr Thr Leu Trp Glu Val Ser Cys Ser Gly Thr Val
325 330 335

Ser Glu Gln Leu Arg Asn Ser Glu Thr Lys Asn Leu Cys Leu Leu Glu
340 345 350

Thr Ala Gln
355

35 (166) INFORMATION FOR SEQ ID NO:165:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1089 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:

```
ATGGGCAACC ACACGTGGGA GGGCTGCCAC GTGGACTCGC GCGTGGACCA CCTCTTTCCG      60
CCATCCCTCT ACATCTTTGT CATCGGCGTG GGGCTGCCCA CCAACTGCCT GGCTCTGTGG      120
GCGGCCTACC GCCAGGTGCA ACAGCGCAAC GAGCTGGGCG TCTACCTGAT GAACCTCAGC      180
ATCGCCGACC TGCTGTACAT CTGCACGCTG CCGCTGTGGG TGGACTACTT CCTGCACCAC      240
10 GACAACTGGA TCCACGGCCC CGGGTCCTGC AAGCTCTTTG GGTTCATCTT CTACACCAAT      300
ATCTACATCA GCATCGCCTT CCTGTGCTGC ATCTCGGTGG ACCGCTACCT GGCTGTGGCC      360
CACCCACTCC GCTTCGCCCC CCTGCGCCGC GTCAAGACCG CCGTGGCCGT GAGCTCCGTG      420
GTCTGGGCCA CGGAGCTGGG CGCCAACTCG GCGCCCCTGT TCCATGACGA GCTCTTCCGA      480
GACCGCTACA ACCACACCTT CTGCTTTGAG AAGTTCCCCA TGGAAGGCTG GGTGGCCTGG      540
15 ATGAACCTCT ATCGGGTGTT CGTGGGCTTC CTCTTCCCGT GGGCGCTCAT GCTGCTGTCTG      600
TACCGGGGCA TCCTGCGGGC CGTGCGGGGC AGCGTGTCCA CCGAGCGCCA GGAGAAGGCC      660
AAGATCGCGC GGCTGGCCCT CAGCCTCATC GCCATCGTGC TGGTCTGCTT TGCGCCCTAT      720
CACGTGCTCT TGCTGTCCCG CAGCGCCATC TACCTGGGCC GCCCTGGGA CTGCGGCTTC      780
GAGGAGCGCG TCTTTTCTGC ATACCACAGC TCACTGGCTT TCACCAGCCT CAACTGTGTG      840
20 GCGGACCCCA TCCTCTACTG CCTGGTCAAC GAGGGCGCCC GCAGCGATGT GGCCAAGGCC      900
CTGCACAACC TGCTCCGCTT TCTGGCCAGC GACAAGCCCC AGGAGATGGC CAATGCCTCG      960
CTCACCTTGG AGACCCCACT CACCTCCAAG AGGAACAGCA CAGCCAAAGC CATGACTGGC      1020
AGCTGGGCGG CCACTCCGCC TTCCAGGGG GACCAGGTGC AGCTGAAGAT GCTGCCGCCA      1080
GCACAATGA                                     1089
```

25 (167) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 362 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
30 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:

	Met	Gly	Asn	His	Thr	Trp	Glu	Gly	Cys	His	Val	Asp	Ser	Arg	Val	Asp	1	5	10	15
5	His	Leu	Phe	Pro	Pro	Ser	Leu	Tyr	Ile	Phe	Val	Ile	Gly	Val	Gly	Leu	20	25	30	
	Pro	Thr	Asn	Cys	Leu	Ala	Leu	Trp	Ala	Ala	Tyr	Arg	Gln	Val	Gln	Gln	35	40	45	
	Arg	Asn	Glu	Leu	Gly	Val	Tyr	Leu	Met	Asn	Leu	Ser	Ile	Ala	Asp	Leu	50	55	60	
10	Leu	Tyr	Ile	Cys	Thr	Leu	Pro	Leu	Trp	Val	Asp	Tyr	Phe	Leu	His	His	65	70	75	80
	Asp	Asn	Trp	Ile	His	Gly	Pro	Gly	Ser	Cys	Lys	Leu	Phe	Gly	Phe	Ile	85	90	95	
15	Phe	Tyr	Thr	Asn	Ile	Tyr	Ile	Ser	Ile	Ala	Phe	Leu	Cys	Cys	Ile	Ser	100	105	110	
	Val	Asp	Arg	Tyr	Leu	Ala	Val	Ala	His	Pro	Leu	Arg	Phe	Ala	Arg	Leu	115	120	125	
	Arg	Arg	Val	Lys	Thr	Ala	Val	Ala	Val	Ser	Ser	Val	Val	Trp	Ala	Thr	130	135	140	
20	Glu	Leu	Gly	Ala	Asn	Ser	Ala	Pro	Leu	Phe	His	Asp	Glu	Leu	Phe	Arg	145	150	155	160
	Asp	Arg	Tyr	Asn	His	Thr	Phe	Cys	Phe	Glu	Lys	Phe	Pro	Met	Glu	Gly	165	170	175	
25	Trp	Val	Ala	Trp	Met	Asn	Leu	Tyr	Arg	Val	Phe	Val	Gly	Phe	Leu	Phe	180	185	190	
	Pro	Trp	Ala	Leu	Met	Leu	Leu	Ser	Tyr	Arg	Gly	Ile	Leu	Arg	Ala	Val	195	200	205	
	Arg	Gly	Ser	Val	Ser	Thr	Glu	Arg	Gln	Glu	Lys	Ala	Lys	Ile	Ala	Arg	210	215	220	
30	Leu	Ala	Leu	Ser	Leu	Ile	Ala	Ile	Val	Leu	Val	Cys	Phe	Ala	Pro	Tyr	225	230	235	240
	His	Val	Leu	Leu	Leu	Ser	Arg	Ser	Ala	Ile	Tyr	Leu	Gly	Arg	Pro	Trp	245	250	255	
35	Asp	Cys	Gly	Phe	Glu	Glu	Arg	Val	Phe	Ser	Ala	Tyr	His	Ser	Ser	Leu	260	265	270	
	Ala	Phe	Thr	Ser	Leu	Asn	Cys	Val	Ala	Asp	Pro	Ile	Leu	Tyr	Cys	Leu				

116

275 280 285

Val Asn Glu Gly Ala Arg Ser Asp Val Ala Lys Ala Leu His Asn Leu
290 295 300

5 Leu Arg Phe Leu Ala Ser Asp Lys Pro Gln Glu Met Ala Asn Ala Ser
305 310 315 320

Leu Thr Leu Glu Thr Pro Leu Thr Ser Lys Arg Asn Ser Thr Ala Lys
325 330 335

Ala Met Thr Gly Ser Trp Ala Ala Thr Pro Pro Ser Gln Gly Asp Gln
340 345 350

10 Val Gln Leu Lys Met Leu Pro Pro Ala Gln
355 360

(168) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 1002 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:

20 ATGGAGTCCT CAGGCAACCC AGAGAGCACC ACCTTTTTTT ACTATGACCT TCAGAGCCAG 60
CCGTGTGAGA ACCAGGCCTG GGTCTTTGCT ACCCTCGCCA CCACTGTCCT GTACTGCCTG 120
GTGTTTCTCC TCAGCCTAGT GGGCAACAGC CTGGTCCTGT GGGTCCTGGT GAAGTATGAG 180
AGCCTGGAGT CCCTCACCAA CATCTTCATC CTCAACCTGT GCCTCTCAGA CCTGGTGTTC 240
GCCTGCTTGT TGCCTGTGTG GATCTCCCCA TACCACTGGG GCTGGGTGCT GGGAGACTTC 300

25 CTCTGCAAAC TCCTCAATAT GATCTTCTCC ATCAGCCTCT ACAGCAGCAT CTTCTTCCTG 360
ACCATCATGA CCATCCACCG CTACCTGTCG GTAGTGAGCC CCCTCTCCAC CCTGCGCGTC 420
CCCACCCTCC GCTGCCGGGT GCTGGTGACC ATGGCTGTGT GGGTAGCCAG CATCCTGTCC 480
TCCATCCTCG ACACCATCTT CCACAAGGTG CTTTCTTCGG GCTGTGATTA TTCCGAACTC 540
ACGTGGTACC TCACCTCCGT CTACCAGCAC AACCTCTTCT TCCTGCTGTC CCTGGGGATT 600

30 ATCCTGTTCT GCTACGTGGA GATCCTCAGG ACCCTGTTCC GCTCACGCTC CAAGCGGCGC 660
CACCGCACGA AAAAGCTCAT CTTGCGCATC GTGGTGGCCT ACTTCCTCAG CTGGGGTCCC 720
TACAACTTCA CCCTGTTTCT GCAGACGCTG TTTCGGACCC AGATCATCCG GAGCTGCGAG 780

GCCAAACAGC AGCTAGAATA CGCCCTGCTC ATCTGCCGCA ACCTCGCCTT CTCCCCTGTC 840
 TGCTTTAACC CGGTGCTCTA TGTCTTCGTG GGGGTCAAGT TCCGCACACA CCTGAAACAT 900
 GTTCTCCGGC AGTTCTGGTT CTGCCGGCTG CAGGCACCCA GCCCAGCCTC GATCCCCCAG 960
 TCCCCTGGTG CCTTCGCCTA TGAGGGCGCC TCCTTCTACT GA 1002

5 (169) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: not relevant

10

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:168:

Met Glu Ser Ser Gly Asn Pro Glu Ser Thr Thr Phe Phe Tyr Tyr Asp
 1 5 10 15
 15 Leu Gln Ser Gln Pro Cys Glu Asn Gln Ala Trp Val Phe Ala Thr Leu
 20 25 30
 Ala Thr Thr Val Leu Tyr Cys Leu Val Phe Leu Leu Ser Leu Val Gly
 35 40 45
 20 Asn Ser Leu Val Leu Trp Val Leu Val Lys Tyr Glu Ser Leu Glu Ser
 50 55 60
 Leu Thr Asn Ile Phe Ile Leu Asn Leu Cys Leu Ser Asp Leu Val Phe
 65 70 75 80
 Ala Cys Leu Leu Pro Val Trp Ile Ser Pro Tyr His Trp Gly Trp Val
 85 90 95
 25 Leu Gly Asp Phe Leu Cys Lys Leu Leu Asn Met Ile Phe Ser Ile Ser
 100 105 110
 Leu Tyr Ser Ser Ile Phe Phe Leu Thr Ile Met Thr Ile His Arg Tyr
 115 120 125
 30 Leu Ser Val Val Ser Pro Leu Ser Thr Leu Arg Val Pro Thr Leu Arg
 130 135 140
 Cys Arg Val Leu Val Thr Met Ala Val Trp Val Ala Ser Ile Leu Ser
 145 150 155 160
 Ser Ile Leu Asp Thr Ile Phe His Lys Val Leu Ser Ser Gly Cys Asp
 165 170 175
 35 Tyr Ser Glu Leu Thr Trp Tyr Leu Thr Ser Val Tyr Gln His Asn Leu
 180 185 190

118

Phe Phe Leu Leu Ser Leu Gly Ile Ile Leu Phe Cys Tyr Val Glu Ile
 195 200 205
 Leu Arg Thr Leu Phe Arg Ser Arg Ser Lys Arg Arg His Arg Thr Lys
 210 215 220
 5 Lys Leu Ile Phe Ala Ile Val Val Ala Tyr Phe Leu Ser Trp Gly Pro
 225 230 235 240
 Tyr Asn Phe Thr Leu Phe Leu Gln Thr Leu Phe Arg Thr Gln Ile Ile
 245 250 255
 10 Arg Ser Cys Glu Ala Lys Gln Gln Leu Glu Tyr Ala Leu Leu Ile Cys
 260 265 270
 Arg Asn Leu Ala Phe Ser His Cys Cys Phe Asn Pro Val Leu Tyr Val
 275 280 285
 Phe Val Gly Val Lys Phe Arg Thr His Leu Lys His Val Leu Arg Gln
 290 295 300
 15 Phe Trp Phe Cys Arg Leu Gln Ala Pro Ser Pro Ala Ser Ile Pro His
 305 310 315 320
 Ser Pro Gly Ala Phe Ala Tyr Glu Gly Ala Ser Phe Tyr
 325 330

(170) INFORMATION FOR SEQ ID NO:169:

20 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 987 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

ATGGACAACG CCTCGTTCTC GGAGCCCTGG CCCGCCAACG CATCGGGCCC GGACCCGGCG 60
 CTGAGCTGCT CCAACGCGTC GACTCTGGCG CCGCTGCCGG CGCCGCTGGC GGTGGCTGTA 120
 CCAGTTGTCT ACGCGGTGAT CTGCGCCGTG GGTCTGGCGG GCAACTCCGC CGTGCTGTAC 180
 30 GTGTTGCTGC GGGCGCCCCG CATGAAGACC GTCACCAACC TGTTCATCCT CAACCTGGCC 240
 ATCGCCGACG AGCTCTTCAC GCTGGTGCTG CCCATCAACA TCGCCGACTT CCTGCTGCGG 300
 CAGTGGCCCT TCGGGGAGCT CATGTGCAAG CTCATCGTGG CTATCGACCA GTACAACACC 360
 TTCTCCAGCC TCTACTTCCT CACCGTCATG AGCGCCGACC GCTACCTGGT GGTGTTGGCC 420
 ACTGCGGAGT CGCGCCGGGT GGCCGGCCGC ACCTACAGCG CCGCGCGCGC GGTGAGCCTG 480

GCCGTGTGGG GGATCGTCAC ACTCGTCGTG CTGCCCTTCG CAGTCTTCGC CCGGCTAGAC 540
 GACGAGCAGG GCCGGCGCCA GTGCGTGCTA GTCTTTCCGC AGCCCGAGGC CTTCTGGTGG 600
 CGCGCGAGCC GCCTCTACAC GCTCGTGCTG GGCTTCGCCA TCCCCGTGTC CACCATCTGT 660
 GTCCTCTATA CCACCCTGCT GTGCCGGCTG CATGCCATGC GGCTGGACAG CCACGCCAAG 720
 5 GCCCTGGAGC GCGCCAAGAA GCGGGTGAAG TTCCTGGTGG TGGCAATCCT GCGGGTGTGC 780
 CTCCTCTGCT GGACGCCCTA CCACCTGAGC ACCGTGGTGG CGCTCACCAC CGACCTCCCG 840
 CAGACGCCGC TGGTCATCGC TATCTCCTAC TTCATCACCA GCCTGACGTA CGCCAACAGC 900
 TGCCTCAACC CCTTCCTCTA CGCCTTCCTG GACGCCAGCT TCCGCAGGAA CCTCCGCCAG 960
 CTGATAACTT GCCGCGCGGC AGCCTGA 987

10 (171) INFORMATION FOR SEQ ID NO:170:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 328 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 15 (D) TOPOLOGY: not relevant

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

Met Asp Asn Ala Ser Phe Ser Glu Pro Trp Pro Ala Asn Ala Ser Gly
 1 5 10 15
 20 Pro Asp Pro Ala Leu Ser Cys Ser Asn Ala Ser Thr Leu Ala Pro Leu
 20 25 30
 Pro Ala Pro Leu Ala Val Ala Val Pro Val Val Tyr Ala Val Ile Cys
 35 40 45
 25 Ala Val Gly Leu Ala Gly Asn Ser Ala Val Leu Tyr Val Leu Leu Arg
 50 55 60
 Ala Pro Arg Met Lys Thr Val Thr Asn Leu Phe Ile Leu Asn Leu Ala
 65 70 75 80
 Ile Ala Asp Glu Leu Phe Thr Leu Val Leu Pro Ile Asn Ile Ala Asp
 85 90 95
 30 Phe Leu Leu Arg Gln Trp Pro Phe Gly Glu Leu Met Cys Lys Leu Ile
 100 105 110
 Val Ala Ile Asp Gln Tyr Asn Thr Phe Ser Ser Leu Tyr Phe Leu Thr
 115 120 125
 Val Met Ser Ala Asp Arg Tyr Leu Val Val Leu Ala Thr Ala Glu Ser

120

	130	135	140
	Arg Arg Val Ala Gly	Arg Thr Tyr Ser Ala	Ala Arg Ala Val Ser Leu
	145	150	155 160
5	Ala Val Trp Gly Ile Val Thr Leu Val Val Leu Pro Phe Ala Val Phe		
		165 170	175
	Ala Arg Leu Asp Asp Glu Gln Gly Arg Arg Gln Cys Val Leu Val Phe		
		180 185	190
	Pro Gln Pro Glu Ala Phe Trp Trp Arg Ala Ser Arg Leu Tyr Thr Leu		
		195 200	205
10	Val Leu Gly Phe Ala Ile Pro Val Ser Thr Ile Cys Val Leu Tyr Thr		
		210 215	220
	Thr Leu Leu Cys Arg Leu His Ala Met Arg Leu Asp Ser His Ala Lys		
		225 230	235 240
15	Ala Leu Glu Arg Ala Lys Lys Arg Val Lys Phe Leu Val Val Ala Ile		
		245 250	255
	Leu Ala Val Cys Leu Leu Cys Trp Thr Pro Tyr His Leu Ser Thr Val		
		260 265	270
	Val Ala Leu Thr Thr Asp Leu Pro Gln Thr Pro Leu Val Ile Ala Ile		
		275 280	285
20	Ser Tyr Phe Ile Thr Ser Leu Thr Tyr Ala Asn Ser Cys Leu Asn Pro		
		290 295	300
	Phe Leu Tyr Ala Phe Leu Asp Ala Ser Phe Arg Arg Asn Leu Arg Gln		
		305 310	315 320
25	Leu Ile Thr Cys Arg Ala Ala Ala		
		325	

(172) INFORMATION FOR SEQ ID NO:171:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1002 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

	ATGCAGGCCG CTGGGCACCC AGAGCCCCTT GACAGCAGGG GCTCCTTCTC CCTCCCCACG	60
35	ATGGGTGCCA ACGTCTCTCA GGACAATGGC ACTGGCCACA ATGCCACCTT CTCCGAGCCA	120
	CTGCCGTTCC TCTATGTGCT CCTGCCCGCC GTGTACTCCG GGATCTGTGC TGTGGGGCTG	180

ACTGGCAACA CGGCCGTCAT CCTTGTAATC CTAAGGGCGC CCAAGATGAA GACGGTGACC 240
 AACGTGTTCA TCCTGAACCT GGCCGTCGCC GACGGGCTCT TCACGCTGGT ACTGCCTGTC 300
 AACATCGCGG AGCACCTGCT GCAGTACTGG CCCTTCGGGG AGCTGCTCTG CAAGCTGGTG 360
 CTGGCCGTCG ACCACTACAA CATCTTCTCC AGCATCTACT TCCTAGCCGT GATGAGCGTG 420
 5 GACCGATACC TGGTGGTGCT GGCCACCGTG AGGTCCCGCC ACATGCCCTG GCGCACCTAC 480
 CGGGGGGCGA AGGTCGCCAG CCTGTGTGTC TGGCTGGGCG TCACGGTCCT GGTTCGTGCC 540
 TTCTTCTCTT TCGCTGGCGT CTACAGCAAC GAGCTGCAGG TCCAAGCTG TGGGCTGAGC 600
 TTCCCGTGGC CCGAGCAGGT CTGGTTCAAG GCCAGCCGTG TCTACACGTT GGTCTGGGC 660
 TTCGTGCTGC CCGTGTGCAC CATCTGTGTG CTCTACACAG ACCTCCTGCG CAGGCTGCGG 720
 10 GCCGTGCGGC TCCGCTCTGG AGCCAAGGCT CTAGGCAAGG CCAGGCGGAA GGTGAAAGTC 780
 CTGGTCTCTG TCGTGCTGGC CGTGTGCCTC CTCTGCTGGA CGCCCTTCCA CCTGGCCTCT 840
 GTCGTGGCCC TGACCACGGA CCTGCCCCAG ACCCACTGG TCATCAGTAT GTCCTACGTC 900
 ATCACCAGCC TCACGTACGC CAACTCGTGC CTGAACCCCT TCCTCTACGC CTTTCTAGAT 960
 GACAACTTCC GGAAGAACTT CCGCAGCATA TTGCGGTGCT GA 1002

15 (173) INFORMATION FOR SEQ ID NO:172:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 333 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: not relevant

20

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

Met Gln Ala Ala Gly His Pro Glu Pro Leu Asp Ser Arg Gly Ser Phe
 1 5 10 15
 25 Ser Leu Pro Thr Met Gly Ala Asn Val Ser Gln Asp Asn Gly Thr Gly
 20 25 30
 His Asn Ala Thr Phe Ser Glu Pro Leu Pro Phe Leu Tyr Val Leu Leu
 35 40 45
 30 Pro Ala Val Tyr Ser Gly Ile Cys Ala Val Gly Leu Thr Gly Asn Thr
 50 55 60
 Ala Val Ile Leu Val Ile Leu Arg Ala Pro Lys Met Lys Thr Val Thr
 65 70 75 80

